

SEARCH REQUEST FORM

Requestor's Name: _____ Serial Number: _____
Date: _____, Phone: _____ Art Unit: _____

Search Topic:

Please write a detailed statement of search topic. Describe specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples or relevant citations, authors, keywords, etc., if known. For sequences, please attach a copy of the sequence. You may include a copy of the broadest and/or most relevant claim(s).

STAFF USE ONLY

Date completed: 07-15-02

Searcher: Beverly E 4997

Terminal time: 20

Elapsed time: _____

CPU time: _____

Total time: 25

Number of Searches: _____

Number of Databases: 1

Search Site

_____ STIC

_____ CM-1

_____ Pre-S

Type of Search

_____ N.A. Sequence

_____ A.A. Sequence

_____ Structure

_____ Bibliographic

Vendors

_____ IG

_____ STN

_____ Dialog

_____ APS

_____ Geninfo

_____ SDC

_____ DARC/Questel

_____ Other CGN

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 12:16:23 ; Search time 58.93 Seconds
(without alignments)
1053.627 Million cell updates/sec

Title: US-09-298-924-6
Perfect score: 2978
Sequence: 1 MTFAYKIDGNEVIFTWAPY.....POHIEGKYFDKGFALYKL 559

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802.*
1: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2978	100.0	559	17 AAR90619	Sulfolobus solfata
2	1748.5	58.7	556	17 AAR92755	Trehalose-releasin
3	1748.5	58.7	556	17 AAR90620	Trehalose-releasin
4	1089.5	36.6	595	22 AAG92072	Sulfolobus acidoca
5	1089.5	36.6	610	22 AAR80206	Corynebacterium gl
6	1060.5	35.6	598	16 AAR77471	Trehalose releasin
7	1058	35.5	597	16 AAR80290	Trehalose releasin
8	1025	34.4	589	16 AAR80289	Trehalose releasin
9	1025	34.4	596	16 AAR77470	Trehalose releasin
10	1024.5	34.4	575	21 AAY85157	Trehalose-releasin
11	1024.5	34.4	575	21 AAY85165	Trehalose-releasin

12	399	13.4	652	16	AAR80037	Bacillus stearothe
13	399	13.4	652	17	AAR96109	Starch-branching-e
14	385.5	12.9	648	10	AAP94635	B. thuringiensis a
15	374.5	12.6	639	12	AAR11271	B. stearothermophil
16	371.5	12.5	718	22	AAB60903	Propionibacterium
17	361	12.1	630	22	AAB69074	Aquifex aeolicus v
18	348	11.7	670	22	AAU60913	Propionibacterium
19	348	11.7	765	20	AAU00870	S. tuberosum isoam
20	346.5	11.6	658	13	AAR23787	Heat-resistant pul
21	345.5	11.6	666	20	AAV34991	C. pneumoniae prot
22	342.5	11.5	793	20	AAU00869	S. tuberosum isoam
23	340	11.4	731	22	AAG91100	C glutamicum prote
24	340	11.4	731	22	AAB79423	Corynebacterium gl
25	339	11.4	764	21	AAV50819	Wheat isoamylase p
26	334	11.2	606	18	AAW36602	S. tuberosum debra
27	333	11.2	1250	21	AAV91279	Group B Streptococ
28	331.5	11.1	618	20	AAW37184	Protein involved i
29	331.5	11.1	798	20	AAW73552	Pullulanase protel
30	331.5	11.1	1938	20	AAW73553	Full length Pullul
31	329.5	11.1	818	20	AAW17523	zea mays SUI starc
32	329.5	11.1	893	17	AAW09257	Bacillus alkaline
33	329.5	11.1	1938	17	AAW09255	Bacillus alkaline
34	315	10.6	772	19	AAW49871	Thermotoga maritim
35	312.5	10.5	621	21	AAW19285	A polypeptide with
36	312	10.5	772	18	AAW34567	Thermotoga maritim
37	309.5	10.4	722	20	AAV35095	C. pneumoniae prot
38	303.5	10.2	829	22	AAE05691	Bacillus acidopull
39	303.5	10.2	862	21	AAV78513	Pullulanase (pulB)
40	303	10.2	931	20	AAV27357	Group B Streptococ
41	300	10.1	498	10	AAV96105	Sequence of amy C
42	299	10.0	928	22	AAE05719	Bacillus deramific
43	296	9.9	762	21	AAW90977	N. denitrificans a
44	295	9.9	555	14	AAW43340	Alpha-glucosidase
45	294.5	9.9	785	19	AAW70886	Protein encoded by

ALIGNMENTS

RESULT 1
ID AAR90619
ID AAR90619 standard; Protein; 559 AA.
XX AC AAR90619;
DT 29-JUN-1996 (first entry)
XX DE Sulfolobus solfataricus amylase for alpha, alpha-trehalose prodn.
XX KW transferase; amylase; Sulfolobus; production; alpha, alpha-trehalose;
XX KW malto-oligosaccharide; hydrolysis.
XX OS Sulfolobus solfataricus.
XX PN W09534642-A.
XX PD 21-DEC-1995.
XX PF 14-JUN-1995; 95WO-JP01189.
XX PR 21-APR-1995; 95JP-0120673.
XX PR 15-JUN-1994; 94JP-0133354.
XX PR 18-AUG-1994; 94JP-0194223.
XX PR 31-OCT-1994; 94JP-0290394.
XX PR 21-NOV-1994; 94JP-0286917.
XX PR 21-NOV-1994; 94JP-0311185.
XX PA (KIRI) KIRIN BEER KK.
XX WPI; 1996-049671/05.
XX DR N-PSDB; AAT12325.
XX PT Sulfolobus spp. derived transferase and amylase - for production of

PT alpha, alpha-trehalose from malto-oligosaccharide(s)

PS Claim 108; Page 235-240; 357pp; Japanese.

XX The amylase is derived from Sulfolobus solfataricus. The amylase acts
XX on a saccharide having at least three sugar units, which are pref.
XX glucose units at the reducing end (the linkage between the first and
XX second glucose units is alpha-1, alpha-1, while the linkage between the
XX second and third glucose units is alpha-1,4), to hydrolyse alpha-1,4
XX linkages within the sugar chain, yielding alpha, alpha-trehalose and also
XX mono- and disaccharide hydrolysis products. The amylase has a mol. wt. of
XX 61 to 64 kDa. It is characterised by working at pH 4.5-5.5 and at 60-85
XX deg.C. It has an isoelectric point of 4.3-5.4 and retains at least
XX 100 percent activity after 6 hrs. at 80 deg.C. It is completely inhibited
XX by 5 mM copper sulphate. Use of a transferase and the amylase in
XX succession on suitable substrates such as malto-oligosaccharides, is
XX useful for the production of alpha, alpha-trehalose.

XX Sequence 559 AA;

Query Match 100.0%; Score 2978; DB 17; Length 559;
Best Local Similarity 100.0%; Pred. No. 3.6e-236;
Matches 559; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MTFAYKIDGNEVIFTLWAPYQKSVKLVLEKGLYMERDEKGYFTITLNNKVRDRYKYV 60

Db 1 mtfaykidgneviftlwapyqskvlkylekgllymerdekgyftitlnnkvrdrkyv 60

Qy 61 LDDASEIPDPASRYQPEGVHGPSQIIQESKEFNNETFLKEDLIIEIHVGFTPEGTFF 120

Db 61 lddaseipdpasryqpegvhpqsglqeskefnnetflkediieihvgftpegtffe 120

Qy 121 GVIRKLDYLKDLGITAIEIMPTAOPGKRDGVDGYLYAVQNSVGGPEGRKLVDEAHK 180

Db 121 gvirkldylkdlgtaitaimptapgkrdgdgydyllyavqnsyggpegfrklvdeahk 180

Qy 181 KGLGVILDVYVNHVGPENYMKLGPYFSQYKTPWGTLTFNFDASDEVRKFIENVEY 240

Db 181 kglgvildvvyvnhvgpegnymklgpyfsqyktpwgtltnfddasdevrkfienvvey 240

Qy 241 WIKENYVDGFRDLDAVHAIDTSPKHILEIADVHKYNRIVIAESDLNDPRVNPKECG 300

Db 241 wikenyvdfgrldavhaidtspkhilleiadvhkynriviaesdlndprvnpkecg 300

Qy 301 YNIDAQWDDFHHSHAYLTGEROGYYTDFGNLDDIVKSYKDVYDGKYNFRKRTHGE 360

Db 301 ynidaqwddfhshihayltgerogytcdfgnlddivksykdvydgkynfrkrthge 360

Qy 361 PVGELDGNFVYIIONHDQVGNRGKGERIILKVDRESYKIAAALYLLSPYIPMIFMGEY 420

Db 361 pvgelsgcnfvvyionhdqvgngkgeriiklvdresykiaaalyllsppyipmifmgey 420

Qy 421 GEENPFYFSDPSKLLIQVREGKKGQDTPDQDSTFNASKLWKIDEEIFSFKI 480

Db 421 geenpfyfsdpsklliqvregkkengqtdpddstfnasklwkideeifsyki 480

Qy 481 LIKMKELSIACDRNVVNGENWLLIKGREYFSLYVFSKSIIEVKYSGTLLSSNNSFP 540

Db 481 likmkelsiacdrnvvnngenwllikgreyslyvfskssievkysgtllssnnsfp 540

Qy 541 QHIEEGKYFDFKGFALYKL 559

Db 541 qhieegkyfcdkgfalykl 559

RESULT 2

AAR92755

ID AAR92755 standard; Protein; 556 AA.

XX

AC AAR92755;

XX

DT 03-AUG-1996 (first entry)

XX Trehalose-releasing thermostable enzyme.

XX Thermostable enzyme; trehalose; sweetener.

XX Sulfolobus acidocaldarius strain ATCC 33909.

XX A09527131-A.

XX 01-FEB-1996.

XX 21-JUL-1995; 95AU-0027131.

XX 04-JUL-1995; 95JP-0189760.

XX 21-JUL-1994; 94JP-0190180.

XX 11-APR-1995; 95JP-0109128.

XX (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.

XX Kubota M, Mitsuzumi H, Sugimoto T;

XX WPI: 1996-106284/12.

XX N-PSDB; AAT16899.

XX Recombinant thermostable enzyme from Sulfolobus acidocaldarius, -
XX releases trehalose from non-reducing saccharide at temps. exceeding
XX 55 degrees Centigrade

XX Claim 2; Page 53-54; 74pp; English.

XX A thermostable enzyme (AAR92755) of Sulfolobus acidocaldarius ATCC
XX 33909 releases trehalose from non-reducing saccharides having a
XX trehalose structure as an end unit and a degree of polymerisation of
XX at least 3. It has a mol.wt. of 54-64 kDa (SDS-PAGE), a pI of
XX 5.6-6.6 and is substantially not inactivated when incubated in aq.
XX solution (pH 7.0) at 85 deg for 60 min. Recombinant enzyme is
XX obtd. by expression of an isolated DNA fragment (AAT16899) in
XX host cells, pref. Escherichia coli, using e.g. vector paluescript
XX II SK(+). The trehalose is useful as a sweetener.

XX Sequence 556 AA;

Query Match 58.7%; Score 1748.5; DB 17; Length 556;
Best Local Similarity 58.9%; Pred. No. 4.1e-135;
Matches 329; Conservative 87; Mismatches 136; Indels 7; Gaps 5;

Qy 2 TFAYKIDGNEVIFTLWAPYQKSVKLVLEKGLYMERDEKGYFTITLNNKVRDRYKYL 61

Db 3 sfgnieknkgiflwapyvnsvklk-lskklipmekndegfveiddeenitysyil 61

Qy 62 DDASEIPDPASRYQPEGVHGPSQIIQESKEFNNETFLKEDLIIEIHVGFTPEGTFF 121

Db 62 edkreipdpasryqplgvhdksqirtdyqildgkvkiedliihvlgvfgsegnfk 121

Qy 122 VIRKLDYLKDLGITAIEIMPTAOPGKRDGVDGYLYAVQNSVGGPEGRKLVDEAHK 181

Db 122 vieklldylkdlgtitaimptapgkrdgdgydyllyavqntyyggpewelaklvneahr 181

Qy 182 GLGVILDVYVNHVGPENYMKLGPYFSQYKTPWGTLTFNFDASDEVRKFIENVEY 241

Db 182 giavildvvyvnhigpegnylglgpyfsdryktpwgtltnfddrgcdqvrkfilenvey 241

Qy 242 IKEYNVGDFRDLDAVHAIDTSPKHILEIADVHKYNRIVIAESDLNDPRVNPKECGY 301

Db 242 fktfkidgfrldavhaidfndspkhilgeiaekahqlkvfaesdlndpkiv--kddcgy 299

Qy 302 NIDAQWDDFHHSHAYLTGEROGYYTDFGNLDDIVKSYKDVYDGKYNFRKRTHGE 361

Db 300 kidaqwddfhshahafitkekdyvqgfriediektfkdvfydgkysyrgtghap 359

Qy 362 VGELDGCNFFVYIIONHDQVGNRGKGERIILKVDRESYKIAAALYLLSPYIPMIFMGEY 421

Db 362 vgelsgcnfvvyionhdqvgngkgeriiklvdresykiaaalyllsppyipmifmgey 421

Db 360 vgdpprkfvvfiqndqvgngngerlsiltktdtlylmaatlylspylpifmggeyy 419

Qy 422 EENPFYFSDSKLIQGVREGKKGQDTPDQDESTFNASKLSWKIDEIFSYKIL 481

Db 420 etnpffifsdvpvlikvreglrkennqmidpqqseafklskswkideevldyykql 479

Qy 482 IKMKELSIACDRRVNVNGENWLIIGREYFSLYVFSKSIEVKYSGTLLSSNNSFPQ 541

Db 480 inirkryn-nckrvkevrrregncitlmeikigiasfddivinskitgnlligi--gfpk 536

Qy 542 HIEGK-YEFDKGFALYKL 559

Db 537 klkdelikvnrvgvyql 555

RESULT 3

AAR90620

ID AAR90620 standard; Protein; 556 AA.

AC AAR90620;

XX 29-JUN-1996 (first entry)

XX Sulfolobus acidocaldarius amylase for alpha, alpha-trehalose prodn.

XX transferase; amylase; Sulfolobus; production; alpha, alpha-trehalose;

KW malto-oligosaccharide; hydrolysis.

XX Sulfolobus acidocaldarius.

XX WO9534642-A.

PN 21-DEC-1995.

XX 14-JUN-1995; 95WO-JP011189.

XX 21-APR-1995; 95JP-0120673.

PR 15-JUN-1994; 94JP-0133354.

PR 18-AUG-1994; 94JP-0194223.

PR 31-OCT-1994; 94JP-0290394.

PR 21-NOV-1994; 94JP-0286917.

PR 21-NOV-1994; 94JP-0311185.

XX (KIRI) KIRIN BEER KK.

PA WPI; 1996-049671/05.

XX N-PSDB; AAT12326.

DR Sulfolobus spp. derived transferase and amylase - for production of

PT alpha, alpha-trehalose from malto-oligosaccharide(s)

XX Claim 112; Page 244-250; 357pp; Japanese.

XX The amylase is derived from Sulfolobus acidocaldarius. The amylase acts

CC on a saccharide having at least three sugar units, which are pref.

CC glucose units at the reducing end (the linkage between the first and

CC second glucose units is alpha-1, alpha-1, while the linkage between the

CC second and third glucose units is alpha-1,4), to hydrolyse alpha-1,4

CC linkages within the sugar chain, yielding alpha, alpha-trehalose and also

CC mono- and disaccharide hydrolysis products. The amylase has a mol. wt. of

CC 61 to 64 kDa. It is characterised by working at pH 4.5-5.5 and at 60-85

CC deg.C. It has an isoelectric point of 4.3-5.4 and retains at least

CC 100 percent activity after 6 hrs. at 80 deg.C. It is completely inhibited

CC by 5 mM copper sulphate. Use of a transferase and the amylase in

CC succession on suitable substrates such a malto-oligosaccharides, is

CC useful for the production of alpha, alpha-trehalose.

XX Sequence 556 AA;

Query Match 58.7%; Score 1748.5; DB 17; Length 556;

Best Local Similarity 58.9%; Pred. No. 4.1e-135;

Matches 329; Conservative 87; Mismatches 136; Indels 7; Gaps 5;

Qy 2 TFAYKIDGNEVIFTLWAPYQYSVKLVLEKGLYEMERDEKGYFTITLNNVVRDRYKYL 61

Db 3 sfngnieknkgifkfwapvynsvklk-lskklpmekndegffeideiddeenitysyil 61

Qy 62 DDASEIPDASRPOPEGVHGPVSQIIQESKEFNFTFKKDLIIYEIHVGTFTPEGTFEG 121

Db 62 edkreipdasryqplgvhdksqlirtdyqildlgkvikiedliiyelhvgtfsqegnfkg 121

Qy 122 VIRKLDVLDGTAIEIMPTAQPGGRDMGYDGVLYAVQNSYGGPGRKLVDEAHKK 181

Db 122 vieklvldlgtlgtielmpvaqfpgnrdwgydgvfiyavqntyggpbwelaklvneahr 181

Qy 182 GLGVILDVYNHVGPEGNMYKLGPFYSQRYKTPWGLTFNFDDAESDEVKRFLENEYEW 241

Db 182 giavilgvvynhigpegnyllgpyfsdryktpwgltfnfdrgcgvrkfilenveyw 241

Qy 242 IKEYNVDFGLDAVHAIIIDTSPKHIIEIADVHVHKYRIVIAESDLNDPRVVPNKEKCY 301

Db 242 fktfkidglrladvhaifdnspkhlqelaeakhlqkfviaesdlndpkiv--kddcgy 299

Qy 302 NIDAAQWDDFHHSTHAYLTGEROGYVTDGNLDDIVKSYKDVYVYDCKYGNFRKTHGEP 361

Db 300 kidaqwwddthhahafitkekdydqfgriediektfkdvfydgkysrygrthgap 359

Qy 362 VGLDGCNFVYVYQNHQDQVGNRGKGERIIRKLVRESYKIAAALYLLSPYIPMIFMGEEYG 421

Db 360 vgdipprkfvyfignhdqvgngngerlsiltktdtlylmaatlylspylpifmggeyy 419

Qy 422 EENPFYFSDSKLIQGVREGKKGQDTPDQDESTFNASKLSWKIDEIFSYKIL 481

Db 420 etnpffifsdvpvlikvreglrkennqmidpqqseafklskswkideevldyykql 479

Qy 482 IKMKELSIACDRRVNVNGENWLIIGREYFSLYVFSKSIEVKYSGTLLSSNNSFPQ 541

Db 480 inirkryn-nckrvkevrrregncitlmeikigiasfddivinskitgnlligi--gfpk 536

Qy 542 HIEGK-YEFDKGFALYKL 559

Db 537 klkdelikvnrvgvyql 555

RESULT 4

AAG92072

ID AAG92072 standard; Protein; 595 AA.

AC AAG92072;

XX 26-SEP-2001 (first entry)

XX C glutamicum protein fragment SEQ ID NO: 5826.

DE C glutamicum protein fragment SEQ ID NO: 5826.

XX Corynebacterium bacterium; amino acid synthesis; vitamin; saccharide;

KW organic acid synthesis.

XX Corynebacterium glutamicum.

PN EP1108790-A2.

XX 20-JUN-2001.

XX 18-DEC-2000; 2000EP-0127688.

XX 16-DEC-1999; 99JP-0377484.

PR 07-APR-2000; 2000JP-0159162.

PR 03-AUG-2000; 2000JP-0280988.

XX (KYOW) KYOWA HAKKO KOGYO KK.

XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;

PI Tateishi N, Senoh A, Ikeda M, Ozaki A;

XX WPI; 2001-376931/40.


```
PR 03-SEP-1999; 99DE-1042086.
PR 03-SEP-1999; 99DE-1042087.
PR 03-SEP-1999; 99DE-1042088.
PR 03-SEP-1999; 99DE-1042095.
PR 03-SEP-1999; 99DE-1042124.
PR 03-SEP-1999; 99DE-1042129.
PR 09-MAR-2000; 2000US-0187970.
XX
PA (BADI ) BASF AG.
XX
XX Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;
PI
XX
DR WPI; 2001-137957/14.
DR N-PSDB; AAF72325.
XX
XX Nucleic acids from Corynebacterium glutamicum encoding metabolic
PT pathway proteins, useful for producing fine chemicals in
PT microorganisms, including organic acids, nonproteinogenic amino acids,
PT and purine and pyrimidine bases.-
XX
XX
PS Claim 20; Page 1722-1725; 1737pp; English.
XX
XX AAF71753 to AAF72330 encode the Corynebacterium glutamicum metabolic
CC pathway (MP) proteins given in AAB79634 to AAB80211. The C. glutamicum
CC MP nucleic acids are useful for the production of fine chemicals
CC in microorganisms, including organic acids, nonproteinogenic amino
CC acids, purine and pyrimidine bases, nucleosides, nucleotides, lipids,
CC saturated and unsaturated fatty acids, diols, carbohydrates, aromatic
CC compounds, vitamins, cofactors, polyketides and enzymes.
XX
XX Sequence 610 AA;
SQ
Query Match 36.6%; Score 1089.5; DB 22; Length 610;
Best Local Similarity 41.5%; Pred. No. 6.7e-81;
Matches 227; Conservative 93; Mismatches 198; Indels 29; Gaps 11;
QY 14 FTLWAPYQKSVKLVLEKGLYMERDEKGYFTITLNNVKVRDRYKYVLDLDAE---IPD 69
DB 43 fswaplphdvhl-ilngeltpmhktgswraei-aptagdrygfsldgsswsktlpd 100
QY 70 PASRYQPEGVHGVSQIIQESKENNETFLKK--EDLIIEIHVGTFTPEGTFFGVIRKLD 127
DB 101 prstsgpdgvhglsevsddsgldgqgtgrilpgsvlyelhvgtfstdgtfegvvdvklp 160
QY 128 YLKDGLGTAIEMPIAQFPCRKRWGVDGYLYXAVONSYGSGPEGFRKLVDAHKKGLGVIL 187
DB 161 ylrldlgvtalelvpqfggnrnwgydgvllwlvahvagyggpaglklldashdagiavyl 220
QY 188 DVYNNHVGPNGYMWKLGYPFSQIKYTPWGLTFNFDDAESDEVKFKFLENVEYWIKEYNV 247
DB 221 dvynnhfgpdgnyngqfgytsgg-stgwgdvvnghdsdevrnylldaarqwfedfhv 279
QY 248 DGRFLDAVHAIDTSPKHIIEETADVVKHN-----RIVIAESDNDPRVNPKEKCGY 301
DB 280 dgrlrdavhslldrgavysllaqitmaedvsaqtgiprsliaselelndpkfvtsreaggf 339
QY 302 NIDAQVDDPHHSITHAVLTGEROGYXTDFCNLDDIVKSKYDVFYDGKYSNFRKTHGEP 361
DB 340 gldaqvddlhahlvsgsrgnyysdsgvstlaktlrevfhtgnystrygrnhgrp 399
QY 362 VGE--LDGCVFVYIYQNHQDQNGRKGRIKLVDRSYKIAAALYLLSPYIPMFGE 419
DB 400 vhpdiptasrfvtytthdgtgrnraigdrpstltpqgqylkaaiyspsytpmlfmg 459
QY 420 YGENPFYFSDFSKLIQGVREGRRKE-----NQGD--TDPQDSTFNAKLSWKIDE 472
DB 460 fgattpfaffcshtdpelnrltsegrkrefarlgwnaddlpspelestftssklwfta 519
QY 473 E---IFSFKILLKMRKELSIACDR--RVNVVNGENWLLIKGREYPSLYFSSKEVYK 527
DB 520 egrindayqllhrlhrtlfgspnlltlevehgenwlsmangrgrilnafsddttvpl 579
```

```
QY 528 SGTILLS 534
DB 580 ggelis 586
RESULT 6
ID AAR77471
XX AAR77471 standard; Protein; 598 AA.
XX
XX AAR77471;
XX
XX 25-JAN-1996 (first entry)
XX
XX Trehalose releasing enzyme.
XX
XX Enzyme; trehalose; alpha-glucosyltrehalose; alpha-maltosyltrehalose;
KW alpha-maltotriosyltrehalose; alpha-maltotetraosyltrehalose;
KW maltopentaosyltrehalose; sweetener; taste-improving agent;
KW stabiliser; filler; excipient; adjuvant; foodstuffs; cosmetics;
KW pharmaceuticals.
XX
XX Arthrobacter sp. Q36.
XX
XX EP671470-A2.
XX
XX 13-SEP-1995.
XX
XX 07-MAR-1995; 95EP-0301474.
XX
XX 07-MAR-1994; 94JP-0059840.
XX 07-MAR-1994; 94JP-0059834.
XX
XX (HAYB ) HAYASHIBARA SEIBUTSU KAGAKU.
XX
XX Hattori K, Kubota M, Sugimoto T, Tsusaki K;
XX
XX WPI; 1995-312772/41.
XX N-PSDB; AAQ98672.
XX
XX DNA encoding a trehalose releasing enzyme - which releases trehalose
PT from a non-reducing saccharide having a trehalose structure as an
PT end unit.
XX
XX Claim 6; Page 27-29; 45pp; English.
XX
XX This enzyme can be used for the preparation of trehalose with high
CC yields and efficiency from non-reducing saccharides such as
CC alpha-glucosyltrehalose, alpha-maltosyltrehalose,
CC alpha-maltotriosyltrehalose, alpha-maltotetraosyltrehalose and
CC maltopentaosyltrehalose. The trehalose can be used as a sweetener,
CC taste-improving agent, quality-improving agent, stabiliser, filler,
CC excipient or adjuvant in food products cosmetics and pharmaceuticals.
XX
XX Sequence 598 AA;
SQ
Query Match 35.6%; Score 1060.5; DB 16; Length 598;
Best Local Similarity 39.7%; Pred. No. 1.6e-78;
Matches 232; Conservative 92; Mismatches 197; Indels 63; Gaps 16;
QY 1 MTFAYKIDGNEVI-----FTLWAPYQKSVKLVLEKG-LYEMER-----DEKGYFTI-- 46
DB 1 mthtypreaakpvlpgarydvwapnaesvtl--lageryamrraetgedagwtaag 58
QY 47 --TLNNVKVRDRYKYVLD-DASEIPDPASRYQPEGVHGVSQIIQESK-EFNNETFLKKE- 101
DB 59 aptdgnvd----ygyllgdgetplpdrtrirgpdgvhalsrtfdpsayswgddawgrel 114
QY 102 -DLIIEIHVGTFTPEGTFFGVIRKLDYLDGLGTAIEIMPIAQFPCRKRWGVDGYLYA 160
DB 115 qgavieilhgtftpegttleaaagklodylagldvdfiellpvnafngthnwgdydgvqwa 174
QY 161 VQNSYSGPEGFRKLVDEAHKKGLGVILDVVYNNHVGPNGYMWKLGYPFSQIKYTPWGLTF 220
```

Db 175 vheayggpeaygrfvdaaaglgvqdvynhlgpsgnylprfgpylkggegnwtgds 234
Qy 221 NFDAEDEVKFFLENYEYKVENYKEDVDFRLDAVHAIDTSPKHLEIAADVVKYNRI 280
Db 235 nldpgsdhvrryildnlamwldryvdglrlidavhalkderavhiledfgaladqisae 294
Qy 281 V-----IAESDLNDPRVVPNPKKCGYNIDAQWDDFHHSIHAYLTGERQGYTDFGNLD 334
Db 295 vgrplilaesdlndprllyprdvngyglgqwsddfhavhvnvtgettygsdfdsala 354
Qy 335 DIVKSKDVYDGYKSNFRKTHGCEPVGELDGCNF-----VVYIQNDHQVGNRGK 386
Db 355 alakvlrdgffhdgsyfrhrhgrpi-----nfsavhpaalvcsqnhdqnratg 408
Qy 387 ERITKLDVRESYKTAALYLLSPYPMIFMGEEYGEENPFYFSDFSKLIQGVREGK 446
Db 409 drlseqtipygsalaaavltlcpftpmllmgeeygastpwwfftshepelgkataegri 468
Qy 447 KE---NGOD-----TDPODESTFNASKLSWKIDEE-----IFSFKILIKMKRKELSIAC-- 492
Db 469 kefermgwdpavvpdpqpetfrskldwaeaaegdharielyrsltalrrstpdltkl 528
Qy 493 ---DRRVNVNGENWLIIGREYFSLVYFSKSSIEVKYSGTLLL 533
Db 529 gfdetqvafedarwlrfrggvqvlfnfseqpvsldgagtail 572

RESULT 7
AAR80290
ID AAR80290 standard; Protein: 597 AA.
XX
AC AAR80290;
XX
DT 19-JAN-1996 (first entry)
XX
DE Trehalose releasing enzyme.
XX
KW Enzyme; trehalose; alpha-glucosyltrehalose; alpha-maltosyltrehalose;
KW alpha-maltotriosyltrehalose; alpha-maltotetraosyltrehalose;
KW maltopentaosyltrehalose; sweetener; taste-improving agent;
KW stabiliser; filler; excipient; adjuvant; foodstuffs; cosmetics;
KW pharmaceuticals.
XX
OS Arthrobacter sp. Q36.
XX
PN EP671470-A2.
XX
PD 13-SEP-1995.
XX
PF 07-MAR-1995; 95EP-0301474.
XX
PR 07-MAR-1994; 94JP-0059840.
PR 07-MAR-1994; 94JP-0059834.
XX
PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
XX
PI Hattori K, Kubota M, Sugimoto T, Tsusaki K;
XX
DR WPT; 1995-312772/41.
DR N-PSDB; AAQ986670.
XX
XX DNA encoding a trehalose releasing enzyme - which releases trehalose
PT from a non-reducing saccharide having a trehalose structure as an
PT end unit.
XX
XX Claim 3; Page 24-25; 45pp; English.
XX
CC This enzyme can be used for the preparation of trehalose with high
CC yields and efficiency from non-reducing saccharides such as
CC alpha-glucosyltrehalose, alpha-maltosyltrehalose,
CC alpha-maltotriosyltrehalose, alpha-maltotetraosyltrehalose and
CC maltopentaosyltrehalose. The trehalose can be used as a sweetener,

CC taste-improving agent, quality-improving agent, stabiliser, filler,
XX excipient or adjuvant in food products cosmetics and pharmaceuticals.
SQ Sequence 597 AA;
Query Match 35.5%; Score 1058; DB 16; Length 597;
Best Local Similarity 40.5%; Pred. No. 2.5e-78;
Matches 229; Conservative 88; Mismatches 191; Indels 58; Gaps 15;
Qy 14 FTLWAPYQSKVKLLEKG-LYEMER-----DEKGVFTI-----TLNNKVRDVRKYKVLVD 62
Db 18 ydwapnaesvtil--laggeryamrraetgpdagwtaagatdgnvd----ygyild 71
Qy 63 -DASEIDPPASRYOPEGVHGPSQIIQESK-EFNNETFLKKE--DLIIYEIHVGFTPEGT 118
Db 72 gdepldprrtrpdpghalsrtfdpsayswqddawgqrelgavayielhlgftpegt 131
Qy 119 FEGVIRKLDYLKDLGITAIEIMP IAQFPGRKDWGYDYVLYAVONSYYGGPEGFRKLVEA 178
Db 132 leaaagkllylaglvdfiellpvnafngthwgydgwqfahvhdyygpeayqrfvdaa 191
Qy 179 HKKGLGVILDVVYVNHVGPENYMKLGPYFSQKYKTPWGLTFNFDDEADSEVRFKILENV 238
Db 192 haaglgvldvvyvnyhlgpsgnylprfgpylkggegnwtgdsvnldgpgsdhvrryildnl 251
Qy 239 EYWIKEYNVGDFRLDAVHAIDTSPKHLEIAADVVKYNRIV-----IAESDLNDPRV 292
Db 252 amwldryvdglrlidavhalkderavhiledfgaladqisaevgrplilaesdlndprl 311
Qy 293 VNPKEKCGYNIDAQWDDFHHSIHAYLTGERQGYTDFGNLDIVKSKDVYDGYKSN 352
Db 312 lyprdvngyglgqwsddfhavhvnvtgettygsdfdsalaakvlrdgffhdgsyys 371
Qy 353 FRKTHGCEPVGELDGCNF-----VVYIQNDHQVGNRGKGERIILKIDRESYKIAAAL 404
Db 372 frerhgrpi-----nfsavhpaalvcsqnhdqnratgdrfslqtpygsalaaavl 425
Qy 405 YLSPYPMIFMGEEYGEENPFYFSDFSKLIQGVREGKKE---NGOD-----TDPOD 457
Db 426 tltqpfpmllmgeeygastpwwfftshepelgkataegrikefermgwdpavvpdpq 485
Qy 458 ESTFNASKLSWKIDEE-----IFSFKILIKMKRKELSIAC-----DRRVNVNGENWLI 507
Db 486 petfrskldwaeaaegdharielyrsltalrrstpdltklgfdetqvafedarwlr 545
Qy 508 KGREYFSLVYFSKSSIEVKYSGTLLL 533
Db 546 irggvqvlfnfseqpvsldgagtail 571

RESULT 8
AAR80289
ID AAR80289 standard; Protein: 589 AA.
XX
AC AAR80289;
XX
DT 19-JAN-1996 (first entry)
XX
DE Trehalose releasing enzyme.
XX
KW Enzyme; trehalose; alpha-glucosyltrehalose; alpha-maltosyltrehalose;
KW alpha-maltotriosyltrehalose; alpha-maltotetraosyltrehalose;
KW maltopentaosyltrehalose; sweetener; taste-improving agent;
KW stabiliser; filler; excipient; adjuvant; foodstuffs; cosmetics;
KW pharmaceuticals.
XX
OS Rhizobium sp. M11.
XX
PN EP671470-A2.
XX
PD 13-SEP-1995.
XX

```
PF 07-MAR-1995; 95EP-0301474.
XX
PR 07-MAR-1994; 94JP-0059840.
PR 07-MAR-1994; 94JP-0059834.
PA (HAYB ) HAYASHIBARA SEIBUTSU KAGAKU.
XX
PI Hattori K, Kubota M, Sugimoto T, Tsusaki K;
XX
DR WPI: 1995-312772/41.
DR N-PSDB; AAQ98669.
XX
XX DNA encoding a trehalose releasing enzyme - which releases trehalose
PT from a non-reducing saccharide having a trehalose structure as an
PT end unit.
XX
XX Claim 3; Page 21-22; 45pp; English.
XX
XX This enzyme can be used for the preparation of trehalose with high
CC yields and efficiency from non-reducing saccharides such as
CC alpha-glucosyltrehalose, alpha-maltosyltrehalose,
CC alpha-maltotriosyltrehalose, alpha-maltotetrasyltrehalose and
CC maltopentaosyltrehalose. The trehalose can be used as a sweetener,
CC taste-improving agent, quality-improving agent, stabiliser, filler,
CC excipient or adjuvant in food products cosmetics and pharmaceuticals.
XX
SQ Sequence 589 AA;

Query Match 34.4%; Score 1025; DB 16; Length 589;
Best Local Similarity 40.3%; Pred. NO. 1.3e-75;
Matches 229; Conservative 80; Mismatches 197; Indels 62; Gaps 16;

QY 14 FTWAPYQKSVKLKLEKG-LYEMER-----DEKGYFTITLNNVKVRDRYKVVLD-DAS 65
Db 10 fdwapeagvtl--laggeryemrrpgngpadegwtaadaptdgadvdygylldgdel 67
QY 66 EIPDPASRYQEGVHGPGSQIIOESKEFNFTFLKEDL---IIEIHVGFTFTPEGTFEGV 122
Db 68 pldprrtrqpegvhalsrtfdpgahrwqdgwggrelgsgvlyelhgftftpgtldaa 127
QY 123 IRKLDYLKDLGTAIEIMPTAQPPGRKDWGYGVYLYAVONSYGGPGPFKLVDAAHKG 182
Db 128 agkldylaglgidfiellpvnaftngthwgydgwqfahvhegyggaayqrfvdaahaag 187
QY 183 LGVILDVYVNHVGPENYVWKLGPYSQKYKTPWGLTFNFDAAESDEVKRFILNVEYWI 242
Db 188 lgvildvvyvnhlgpsgnylprygyplkhgegnwtgdsvnldgpgsdhvrqyildnvmwl 247
QY 243 KEYNDGFRDLDAVHAIDTSPKHILEE---IADVWHKYN---RIVIAESDLNDPRVVPNK 296
Db 248 rdyrvdglridavhaikderavhileefgaladalsesegrpitliaesdlnprllypr 307
QY 297 EKGYNIDAQWDDFHHSIHAYLTGERQGYTDFGNLDDIVKYNKDVYVDGYKSNFRRK 356
Db 308 dvngyglagqwsddfhahvhnvsgtgyysdfslgalakvlrdgffhdgysysfgr 367
QY 357 THCEPVELDGNF-----VYIQNHQDVGNCRKGERIILVDRESYKIAAALYLLS 408
Db 368 chrpl-----nfsavhpaalvvcsqnhdqignratgdrisqslpygsialaavltg 421
QY 409 PYIPMIFMGEGEENPFYFSDFSKSLQGVREGKKKE---NGQD-----TPQDESTF 461
Db 422 pftpmfmgeyatpwwqftshpepelgkatakagrirefermrgwdpavvpdpqpetf 481
QY 462 NAKLSWK-----IDEIFSYKILTKMRK-----ELSTACDRRVNVNGENWLIKGR 510
Db 482 trskldwaesasgdharllelysltlrrstpelarlga-dtavefdddarwl----- 535
QY 511 EYFS-----LYVFSKSIEVKYSGTLLL 533
Db 536 rywrgvqvlnfadrpisdprgtall 563
```

```
RESULT 9
AAR77470
ID AAR77470 standard; Protein; 596 AA.
XX
XX AAR77470;
XX
XX 25-JAN-1996 (first entry)
XX
XX Trehalose releasing enzyme.
XX
XX Enzyme; trehalose; alpha-glucosyltrehalose; alpha-maltosyltrehalose;
KW alpha-maltotriosyltrehalose; alpha-maltotetrasyltrehalose;
KW maltopentaosyltrehalose; sweetener; taste-improving agent;
KW stabiliser; filler; excipient; adjuvant; foodstuffs; cosmetics;
KW pharmaceuticals.
XX
XX Rhizobium sp. M11.
XX
XX EP671470-A2.
XX
XX 13-SEP-1995.
XX
XX 07-MAR-1995; 95EP-0301474.
XX
XX 07-MAR-1994; 94JP-0059840.
XX
XX 07-MAR-1994; 94JP-0059834.
XX
XX (HAYB ) HAYASHIBARA SEIBUTSU KAGAKU.
XX
XX Hattori K, Kubota M, Sugimoto T, Tsusaki K;
XX
XX WPI: 1995-312772/41.
XX
XX N-PSDB; AAQ98671.
XX
XX DNA encoding a trehalose releasing enzyme - which releases trehalose
PT from a non-reducing saccharide having a trehalose structure as an
PT end unit.
XX
XX Claim 3; Page 27-29; 45pp; English.
XX
XX This enzyme can be used for the preparation of trehalose with high
CC yields and efficiency from non-reducing saccharides such as
CC alpha-glucosyltrehalose, alpha-maltosyltrehalose,
CC alpha-maltotriosyltrehalose, alpha-maltotetrasyltrehalose and
CC maltopentaosyltrehalose. The trehalose can be used as a sweetener,
CC taste-improving agent, quality-improving agent, stabiliser, filler,
CC excipient or adjuvant in food products cosmetics and pharmaceuticals.
XX
SQ Sequence 596 AA;

Query Match 34.4%; Score 1025; DB 16; Length 596;
Best Local Similarity 40.3%; Pred. NO. 1.3e-75;
Matches 229; Conservative 80; Mismatches 197; Indels 62; Gaps 16;

QY 14 FTWAPYQKSVKLKLEKG-LYEMER-----DEKGYFTITLNNVKVRDRYKVVLD-DAS 65
Db 17 fdwapeagvtl--laggeryemrrpgngpadegwtaadaptdgadvdygylldgdel 74
QY 66 EIPDPASRYQEGVHGPGSQIIOESKEFNFTFLKEDL---IIEIHVGFTFTPEGTFEGV 122
Db 75 pldprrtrqpegvhalsrtfdpgahrwqdgwggrelgsgvlyelhgftftpgtldaa 134
QY 123 IRKLDYLKDLGTAIEIMPTAQPPGRKDWGYGVYLYAVONSYGGPGPFKLVDAAHKG 182
Db 135 agkldylaglgidfiellpvnaftngthwgydgwqfahvhegyggaayqrfvdaahaag 194
QY 183 LGVILDVYVNHVGPENYVWKLGPYSQKYKTPWGLTFNFDAAESDEVKRFILNVEYWI 242
Db 195 lgvildvvyvnhlgpsgnylprygyplkhgegnwtgdsvnldgpgsdhvrqyildnvmwl 254
QY 243 KEYNDGFRDLDAVHAIDTSPKHILEE---IADVWHKYN---RIVIAESDLNDPRVVPNK 296
```

Db 255 rdyrvdgrldavhalderavhileefgaladalssesgrrpltliaesdlnmprillypr 314
 Qy 297 EKCQYNDIAQWDDPHHSIHAAYLGERQGYTDFGNLDDIVKSYKDVYDGYKSNFRRK 356
 Db 315 dvngyglagwsddfhavhvnvsgettgyysdfdsigalakvirdgffhdgssysfgr 374
 Qy 357 THGEPVGLDCNF-----VVIYQNHQDVGNRGKGERIKLVDRSYKIAAALYLLS 408
 Db 375 chgrpi-----nfsavhpaalvcvsnhdqignratgdrisqispygsalaaavltlg 428
 Qy 409 PYIMFNGEYGEENFYFSDSKLIQGVREGKKE---NGQD-----TDPQDESTF 461
 Db 429 pftmfmgeeygattpwqfftshepeelgkataegriferefermgwdpavvvpdpqdbetf 488
 Qy 462 NASLWNR-----IDEEFYSFKILMKRK-----ELSIACDRRVNVNGENWLIIGKR 510
 Db 489 trskldwaesagdharrllelyrlsltrstpelarlgfa-dtafevdddarwl----- 542
 Qy 511 EYFS-----LVFSSKSIEVKYSGTLLL 533
 Db 543 rywggvqvlnfadrlsldrptall 570
 RESULT 10
 ID AAY85157 standard; protein; 575 AA.
 XX AC AAY85157;
 XX DT 23-JUN-2000 (first entry)
 XX DE Trehalose-releasing enzyme amino acid sequence.
 XX KW Non-reducing saccharide-forming enzyme; trehalose; disaccharide; food;
 XX KW high yield; trehalose releasing enzyme; moisture retention; cosmetic;
 XX KW production.
 XX OS Arthrobacter sp. S34.
 XX PN EP990704-A2.
 XX PD 05-APR-2000.
 XX PF 13-SEP-1999; 99EP-0307220.
 XX PR 11-SEP-1998; 98JP-0258394.
 XX PR 11-DEC-1998; 98JP-0352252.
 XX PR 26-JAN-1999; 99JP-0016931.
 XX PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
 XX PI Yamamoto T, Maruta K, Kubota M, Fukuda S, Miyake T;
 XX DR WPI; 2000-273225/24.
 XX DR N-PSDB; AAA10503.
 XX PT New mesophilic enzyme from Arthrobacter, used to convert starch
 XX PT hydrolylate into non-reducing sugars, particularly trehalose, useful as
 XX PT a sweetener in foods, pharmaceuticals and cosmetics -
 XX PS Claim 29; Page 47-48; 93pp; English.
 XX CC This sequence represents a trehalose-releasing enzyme amino acid sequence
 CC CC from Arthrobacter sp. S34. This enzyme specifically hydrolyses a
 CC CC non-reducing saccharide having a trehalose structure as an end unit. The
 CC CC invention relates to this enzyme and also to a non-reducing saccharide
 CC CC forming enzyme that forms a non-reducing saccharide having a trehalose
 CC CC structure as an end unit by reducing partial starch hydrolysates. The two
 CC CC enzymes can be used to produce trehalose, a disaccharide consisting of
 CC CC two moles of glucose bound at their reducing residues. The disaccharide
 CC CC is substantially free of reducibility and has moisture retaining
 CC CC abilities. The enzymes are used to produce non-reducing sugars, they are

CC particularly used for further enzymatic conversion of starch hydrolysate
 CC to trehalose. The non-reducing sugars produced by the enzymes are useful
 CC in foods, pharmaceuticals and cosmetics. Non-reducing sugars are
 CC compatible with amino acids and proteins, they do not cause browning, and
 CC have good moisture-retaining properties. The enzymes produce the
 CC non-reducing sugars in high yield and function at a medium temperature
 CC range i.e. between 40 degrees celsius and 60 degrees celsius, and in an
 CC acidic pH range, i.e. a pH of less than 7. Production of non-reducing
 CC sugars using the enzymes eliminates the need to convert a starch
 CC hydrolysate into sugar alcohols (non-reducing) by catalytic
 CC hydrogenation.
 XX SQ Sequence 575 AA;

Query Match 34.4%; Score 1024.5; DB 21; Length 575;
 Best Local Similarity 39.4%; Pred. No. 1.4e-75;
 Matches 215; Conservative 90; Mismatches 189; Indels 51; Gaps 13;

Qy 14 FTLWAPYQKSVKLVLEKGLYE--MERDEKGYFTITL---NNVKVDRYKYVLDDASEIP 68
 Db 5 fpwapqaaqvtl-vvgggraelpltrdengwalgqpwgdgdlvd-ygylvdgkqpf 62
 Qy 69 DPASRYQPEGVHGPSQIIQESKEFNFTFLKED-----LIIYEHVHFTFPBGT 119
 Db 63 dprslrqrgrvh-----elgrefdparyawgddgrldtgaviyellhvgftpegtl 116
 Qy 120 EGVIKLDYLKDLGITAITIEMPIAQPPGKRDWGYVLYAVQNSYGGPGRKLVDEAH 179
 Db 117 dsairldhvirigvdavellpvnafngthgwydgvlywavyahpepygggyayqrfvdach 176
 Qy 180 KGLGVILDVYVNHVGPENYMKLGPYFSQYKTPMGLTFNFDDESDERVKFLENVE 239
 Db 177 arglavqdvvnhlpgpsgnhlpdfigpylgsgaantwgdalndgplsddevrtyldnav 236
 Qy 240 YWKEYNVDFRLDVAHAIIDTSPKHILEIA-----DVVHKYNR--IIVIAESDLNDRPV 293
 Db 237 ywlrmdhadglridavhalrdaralhlleelaarvdelagelgrpltliaesdlnpkli 296
 Qy 294 NPKEKCYNIDAOVWDDFHHSIHAYLGERQGYTDFGNLDDIVKSYKDVYDGYKSNF 353
 Db 297 rsraahgygidagwddvhhavhanvcgetvgyyadfggalvkvfgrgfhdtgwtssf 356
 Qy 354 RRKTHGEPVG-ELDGCNFFVYIYQNHQDVGNRGKGERIKLVDRSYKIAAALYLLSPYIP 412
 Db 357 rerhhgrpldpdipfrllvafaqhdqvgpravgdrmsaqvgegsalaaalvllgft 416
 Qy 413 MIPWGEYGEENFYFSDSKLIQGVREGKKE---NGQD-----TDPQDESTFNASK 465
 Db 417 mlfmgeewgartpwqfftshepeelgeatargriaefarmgwdpavvvpdpqdpattarsh 476
 Qy 466 LSWKIDEE-----IFSPYKILIKMKRELSTACDRRVNVNGENWLIIGKREYSLVYFSK 520
 Db 477 ldwseperphagllafytdlialrlrelpyvdapare-----vdeadeargvfafsr 526
 Qy 521 SSIEV 525
 Db 527 gplrv 531

RESULT 11
 ID AAY85165
 XX AAY85165 standard; Protein; 575 AA.

XX AC AAY85165;
 XX DT 23-JUN-2000 (first entry)
 XX DE Trehalose-releasing enzyme amino acid sequence.
 XX KW Non-reducing saccharide-forming enzyme; trehalose; disaccharide; food;
 XX KW high yield; trehalose releasing enzyme; moisture retention; cosmetic;
 XX KW production.

Db 465 -----faqfde-----wkfedeldwvlf 482

QY 475 -FSFYKILIRKRELSTACDRRVNV-----NGENWLIIRKREYFSLYVF 518

Db 483 dfehrkmndymkel-lacykrykpfyeldhpggfewidvhaeq-sifsf 532

RESULT 14

AAAP94635

ID AAP94635 standard; protein; 648 AA.

AC AAP94635;

XX

XX 29-JUN-1990 (first entry)

DT

XX

DE B. thuringiensis alpha-amylase.

XX

KW Alpha-amylase; silage; probiotic; lactic acid bacteria;

KW Lactobacillus plantarum; E. faecalis;

KW Bacillus thuringiensis strain H1.1 serotype H1 var thuringiensis.

XX

OS Bacillus thuringiensis.

PN WO8901970-A.

XX

XX 09-MAR-1989.

PD

XX

PF 02-SEP-1988; 88WO-EP00813.

PR 02-SEP-1987; 87EP-0401972.

XX

PA (PLAN-) PLANT GENETIC SYST (UYLO-).

XX

XX Michiels F, Delcour J, Mahillon J, Joos H, Platteeuw C, Josson K;

XX WPI; 1989-085544/11.

DR N-PSDB; AAN91246.

XX

PT Transformed lactic acid bacteria -

PT using exogenous DNA high codes enzyme that breaks down

PT oligosaccharide and/or polysaccharide, used in silage or as a

PT probiotic

XX

PS Fig 2; pages 2/35-3/35; 112pp; English.

XX

CC The inventors claim an inoculum for silage which comprises lactic

CC acid bacteria, esp. Lactobacillus plantarum and E. faecalis, transformed

CC with an exogenous DNA encoding an enzyme which breaks down an oligosac-

CC -haride and/or a polysaccharide into a fermentable carbohydrate. The

CC exogenous enzyme is especially an alpha-amylase (AAP94635) encoded by

CC the DNA sequence in AAN91246 (claims 2, 6 and 10). The transformed lactic

CC acid bacteria can break down polysaccharides and oligosaccharides in a

CC silage crop to provide a complete source of the fermentable carbohydrates

CC that the lactic acid bacteria need for producing lactic acid. Also

CC claimed is a probiotic characterised by the transformed lactic acid

CC bacteria which can be used for establishing and maintaining optimal

CC intestinal flora in animals, including humans.

XX

SQ Sequence 648 AA;

Query Match 12.9%; Score 385.5; DB 10; Length 648;

Best Local Similarity 23.7%; Pred. No. 5.3e-23;

Matches 149; Conservative 117; Mismatches 231; Indels 131; Gaps 29;

QY 4 AYKIDNEVI-----FTLMAPYQKSVK-LKVLEKGLYEMER-----DEKGYFTITLN 49

Db 25 synfaghvvtedegvfrvfwphakamsvvgdfnewdyeqhknlgvteegiwalflp 84

QY 50 NKKVRDRYKVVLD-----DASEIPDPASRYQPEGVHGPSQIIQ-ESKEFNNEVF----- 97

Db 85 hieegeiykyaetlagdvilkadpyavayevrpntasvfdikgyewndknwrkkrkkk 144

QY 98 -LKKEDLLIIEIHVGTF--TPEGTF-----EGVIRKLDYLDKLGITAIETIMPAAQFPGKRD 150

Db 145 siykeamtvyelhfgskkkkedgtlysermveelipyyvvehqfthieimplvehpydrs 204

QY 151 WGVGYYLAVQNSYSGPEGFRKLVDEAHKKGLGVILDVYVNVH--GPEGNYMVKLGPF 208

Db 205 wgyqgtgyaatsrfgtphdmbfvdchekygyigvildwvpgfhckdahglylfdgtpy 264

QY 209 SOKYK-----TPWGLTFNFDAESDEVRKFIENVEYWIKEYNVVDGFRDLDAVHAII-- 259

Db 265 qykdqvpqenpvwg-tvnfdlgr-evrnfliisnalifwmryfhiidgrvdavanmlywnk 322

QY 260 ---DTPSKH---ILEETADVVKYNR--IVIAESDLNDPRVNPKEKCGYNDAQWDDF 311

Db 323 egqeqnehavsfrelneavfaededfImtaedstawplvtpttyegglgfnkykw---- 378

QY 312 HHSIHAYLTGEROYYTDFGNLDDIVKSKDVFYDGYKSNPRKTHGEPVCEL---DGC 368

Db 379 -----nmgmwmdvfk-----ymecapeykrkhihekmtfsillyase 414

QY 369 NFVYIQNHDOVGNRGKGERIKLVDRSEYKIAAALYLLSPYI-----PMIPMGEEYGE 422

Db 415 nfilpi-shdevvh-gk-ksllnkmppgydwkfaqlrlygyffthpgeklfmggefgq 471

QY 423 ENPFYFSDFSKLIQGVREGKKENGQDTPQDESTFNASKLSWKID----- 471

Db 472 -----fdewkldiedlwnldhfemhrymhdynelialykrskplwqldhspegfwid 525

QY 472 -----EEIFSYKILIKMRKELSIACDRRVNVVNGENWLI-INKGREYFSLYVFSKSIEV 525

Db 526 annedqsfirgqgkedalvvvcn--ftkatyenykvpgvdpfeyneilnsdae--- 580

QY 526 KYSGT-----LLLSNNSPFOHIE 544

Db 581 qyggsgqvnkkrkltilepynmqahve 608

RESULT 15

AAAR11271

ID AAR11271 standard; Protein; 639 AA.

XX

AC AAR11271;

XX

DT 29-MAY-1991 (first entry)

XX

DE B.steothermophilus branching enzyme.

XX

KW Branching enzyme; amylaceous substances; transferase;

KW thermostability.

XX

OS Bacillus steothermophilus.

XX

PN EP418945-A.

XX

PD 27-MAR-1991.

XX

XX 22-AUG-1990; 90EP-0202264.

XX

PR 23-AUG-1989; 89NL-0002128.

XX

PA (COOP-) COOP VERKOOP AVEBE.

PI Kiel JAKW, Boels JM, Beldman G, Venema G;

XX WPI; 1991-088587/13.

DR N-PSDB; AAR11072.

XX

PT Branching enzyme obtd. from Bacillus steothermophilus - has good

PT thermostability, for use in modifying starch-like substances

XX Disclosure; fig 4; 13pp; English.

CC This branching enzyme is derived from B.steothermophilus and is

CC encoded by the glgB gene. It is stable at high temps. and can be
CC produced in large amts. using recombinant techniques. It is a
CC transferase allowing formation of alpha(1-6) branches in starch-like
CC substances. It is useful for extending the shelf life of such
CC materials and can be used in the prepn. of human and animal foods.
XX

SQ Sequence 639 AA;

```
Query Match          12.6%; Score 374.5; DB 12; Length 639;
Best Local Similarity 23.6%; Pred. No. 4.2e-22;
Matches 141; Conservative 99; Mismatches 190; Indels 167; Gaps 29;

Qy  4 AYKIDGNEVI-----FTLWAPYQKSVKL-----KVLEKGLYEMERD 39
   :|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 22 syelfgahvikngmvgtrfcvapharevrlvgsfnwngtnfnlmlkvsngqvmifip 81
   :|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 40 E-KGYFTITLNNVKVRDRYKYL--DDASEI--PDPAARYQPEGVGPSQIIO-ESKEF 92
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 82 enleghl-----ykeittndgnvllksdpvafyselrphntasivynikgyqw 129
   :|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 93 NNEIFLKK-----EDLIIEIYHVGTF--TPEGTF-----EGVIRKLDYLKLGITAIE 138
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 130 ndqtwrrkkqrlydgpflfiyehfsgwkkkedgsftyqemaeeilpyvlehgfthie 189
   :|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 139 IMPIAQFPKRDWGYDYLYAVQNSYGGPEGFRKLVDEAHKGLGVILDVYVNHV--GP 196
   :|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 190 llplvehpfdrsyqgigysatsrygtphdlmyfidrchqagigvildwvpgfhckds 249
   :|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 197 EGNWVKLGPFY-----SQKYKTPWGLTFNPDDAESDEVKFIENVEYWIKEYNVGFR 251
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 250 hglymfdgapayeyanmqdrenyvwg-tanf-dlgkpevrsflisnalfwmeythvdgfr 307
   :|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 252 LDVAVHAI-----DTSPKHLEIADVHVKY--NRIVIAESDLNDPRVVPKKEC 299
   :|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 308 vdavannmlywnpsdvlykntyavelfqlkinetvfyaydpnlmliaedstdwprvtapydg 367
   :|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 300 GYNIDAQWVDDFHHSIHAYLTGERQGYTDFGNLDDIVKSYKDVYVDGKYSNFRRTTHG 359
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 368 glgfnykw-----nmgwmdil-----tymetppehrkyvhn 399
   :|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 360 EPGELGDNFVYIQN-----HDQVGNRCKGERIINKLVDRSYKTAAL-----YLLS- 408
   :|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 400 kvtfsl----lyaysenfilpfsdhdevvh-gkksllskmpotyeeekfaqlrlygyllth 454
   :|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 409 PYIPMIFMGEGYGEENPFYFPFSDSKLIQGVREGRKKENGQDTPQDESTFNASKLSW 468
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 455 pgkllfmggf9q-----fdewkd-----leqldw 480
   :|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 469 KIDEEIFSFYKILIKMRKELSIACDRRVNVV-----NGENWLIIRKREYFSLYVF 518
   :|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 481 mlfd--fdmhrnmmyvkel-lkcykcykplyeldhspdgfewidvhnraq-sifsf 533
   :|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

Search completed: July 15, 2002, 12:17:32
Job time: 69 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 12:16:28 ; Search time 25.09 Seconds
(without alignments)
544.197 Million cell updates/sec

Title: US-09-298-924-6
Perfect score: 2978
Sequence: 1 MTFAYKIDGNEVFTLWAPY.....PQHIEGKYFDKGFALYKL 559

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*

- 1: /cgn2_6/protdata/2/1aa/5A_COMB.pep:*
- 2: /cgn2_6/protdata/2/1aa/5B_COMB.pep:*
- 3: /cgn2_6/protdata/2/1aa/6A_COMB.pep:*
- 4: /cgn2_6/protdata/2/1aa/6B_COMB.pep:*
- 5: /cgn2_6/protdata/2/1aa/FCUS_COMB.pep:*
- 6: /cgn2_6/protdata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2978	100.0	559	US-09-242-690A-15	Sequence 15, Appl
2	1748.5	58.7	556	US-08-505-377-1	Sequence 1, Appl
3	1748.5	58.7	556	US-08-798-269-1	Sequence 1, Appl
4	1748.5	58.7	556	US-09-055-210-1	Sequence 1, Appl
5	1060.5	35.6	598	US-08-399-646-14	Sequence 14, Appl
6	1060.5	35.6	598	US-08-607-321-14	Sequence 14, Appl
7	1060.5	35.6	598	US-08-961-240-14	Sequence 14, Appl
8	1060.5	35.6	598	US-08-605-501-14	Sequence 14, Appl
9	1059	35.6	597	US-08-399-646-4	Sequence 4, Appl
10	1059	35.6	597	US-08-607-321-4	Sequence 4, Appl
11	1059	35.6	597	US-08-961-240-4	Sequence 4, Appl
12	1059	35.6	597	US-08-605-501-4	Sequence 4, Appl
13	1025	34.4	589	US-08-399-646-2	Sequence 2, Appl
14	1025	34.4	589	US-08-607-321-2	Sequence 2, Appl
15	1025	34.4	589	US-08-961-240-2	Sequence 2, Appl
16	1025	34.4	589	US-08-605-501-2	Sequence 2, Appl
17	1025	34.4	596	US-08-399-646-12	Sequence 12, Appl
18	1025	34.4	596	US-08-607-321-12	Sequence 12, Appl
19	1025	34.4	596	US-08-961-240-12	Sequence 12, Appl
20	1025	34.4	596	US-08-605-501-12	Sequence 12, Appl
21	399	13.4	652	US-08-528-026C-4	Sequence 4, Appl
22	346.5	11.6	713	US-09-346-237-9	Sequence 9, Appl
23	343	11.5	718	US-09-346-237-10	Sequence 10, Appl
24	334	11.2	606	US-09-187-124-2	Sequence 2, Appl
25	331.5	11.1	893	US-09-514-302-4	Sequence 4, Appl
26	331.5	11.1	1938	US-09-514-302-2	Sequence 2, Appl
27	329.5	11.1	818	US-08-410-784A-2	Sequence 2, Appl

28	329.5	11.1	818	US-09-346-237-11	Sequence 11, Appl
29	321.5	10.8	740	US-08-410-784A-4	Sequence 4, Appl
30	303.5	10.2	829	US-09-514-599-6	Sequence 6, Appl
31	303.5	10.2	862	US-09-346-237-1	Sequence 1, Appl
32	295	9.9	555	US-08-039-777-3	Sequence 3, Appl
33	293.5	9.9	555	US-08-611-361A-3	Sequence 3, Appl
34	293.5	9.9	921	US-09-514-599-2	Sequence 2, Appl
35	291	9.8	915	US-09-346-237-2	Sequence 2, Appl
36	291	9.8	928	US-09-514-599-4	Sequence 4, Appl
37	286	9.6	928	US-08-474-140-11	Sequence 11, Appl
38	286	9.6	928	US-08-477-630-11	Sequence 11, Appl
39	286	9.6	928	US-08-472-293-11	Sequence 11, Appl
40	286	9.6	928	US-08-474-545-11	Sequence 11, Appl
41	286	9.6	928	US-08-478-341-11	Sequence 11, Appl
42	286	9.6	928	US-08-996-733-11	Sequence 11, Appl
43	285	9.6	963	US-08-537-002A-3	Sequence 3, Appl
44	285	9.6	963	US-08-863-010-3	Sequence 3, Appl
45	285	9.6	963	US-09-024-429-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1

US-09-242-690A-15

; Sequence 15, Application US/09242690A

; Patent No. 6284534

; GENERAL INFORMATION:

; APPLICANT: KONDO, KEIJI

; APPLICANT: MIURA, YUTAKA

; TITLE OF INVENTION: YEAST VECTOR AND METHOD OF PRODUCING PROTEINS USING THE

; TITLE OF INVENTION: SAME

; FILE REFERENCE: 049441/0118

; CURRENT APPLICATION NUMBER: US/09/242,690A

; PRIOR FILING DATE: 1999-02-23

; PRIOR APPLICATION NUMBER: PCT/JP97/02924

; PRIOR FILING DATE: 1997-08-22

; PRIOR APPLICATION NUMBER: JP 8/241062

; PRIOR FILING DATE: 1996-08-23

; NUMBER OF SEQ ID NOS: 66

; SOFTWARE: Patent In Ver. 2.1

; SEQ ID NO 15

; LENGTH: 559

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: Synthetic protein

US-09-242-690A-15

Query Match 100.0%; Score 2978; DB 4; Length 559;
Best Local Similarity 100.0%; Pred. No. 4.1e-250; Indels 0; Gaps 0;
Matches, 559; Conservative 0; Mismatches 0;

QY	1	MTFAYKIDGNEVFTLWAPYQKSVKLVLEKGLYEMERDEKGYFTITLNNVKVDRYKYV	60
Db	1	MTFAYKIDGNEVFTLWAPYQKSVKLVLEKGLYEMERDEKGYFTITLNNVKVDRYKYV	60
QY	61	LDASEIPDASRYQPEGVHGPSQIIQESKEFNFTFLKKEDLIYIEIHVGTTPGTFEE	120
Db	61	LDASEIPDASRYQPEGVHGPSQIIQESKEFNFTFLKKEDLIYIEIHVGTTPGTFEE	120
QY	121	GVTRKLDYLDLGLGITAIEIMPYQKSVKLVLEKGLYEMERDEKGYFTITLNNVKVDRYKYV	180
Db	121	GVTRKLDYLDLGLGITAIEIMPYQKSVKLVLEKGLYEMERDEKGYFTITLNNVKVDRYKYV	180
QY	181	KGLGVILDVVYVNHVGPENYMWKLGYPFSQYKYPGLTGFNFDDAESDEVKRFLEINVEY	240
Db	181	KGLGVILDVVYVNHVGPENYMWKLGYPFSQYKYPGLTGFNFDDAESDEVKRFLEINVEY	240
QY	241	WIKEYNVDGFRDLDAVHAIDTSPKHILEETADVHHYKNRIVIAESDLNDPRVVPKKECG	300
Db	241	WIKEYNVDGFRDLDAVHAIDTSPKHILEETADVHHYKNRIVIAESDLNDPRVVPKKECG	300

102(e) date for
this patent is 2/23/99
Senior to mine

QY 301 YNIDAQWDDFHHSHAYLTGERQYTTDFGNLDDIVKSKDVFYDGYKSNFRKTHGE 360
Db 301 YNIDAQWDDFHHSHAYLTGERQYTTDFGNLDDIVKSKDVFYDGYKSNFRKTHGE 360
QY 361 PVGELDCGNFVYIQNDHQVGNRGKRIIKLVDRSYKIAAALYLLSPYIPMFMGEEY 420
Db 361 PVGELDCGNFVYIQNDHQVGNRGKRIIKLVDRSYKIAAALYLLSPYIPMFMGEEY 420
QY 421 GEENPFYFSDSKLIQVREGKRGKENGODTDPQDESTFNASKLSWKIDEEIFSFKI 480
Db 421 GEENPFYFSDSKLIQVREGKRGKENGODTDPQDESTFNASKLSWKIDEEIFSFKI 480
QY 481 LMKRKELSIACDRRVNVNGENWLIIGREYFSLYVFSKSSIEVKYSGTLLSSNNSFP 540
Db 481 LMKRKELSIACDRRVNVNGENWLIIGREYFSLYVFSKSSIEVKYSGTLLSSNNSFP 540
QY 541 QHIEGKYFDPKGFALYKL 559
Db 541 QHIEGKYFDPKGFALYKL 559
RESULT 2
US-08-505-377-1
; Sequence 1, Application US/08505377
; Patent No. 5856146
; GENERAL INFORMATION:
; APPLICANT: MITSUZUMI, Hitoshi
; APPLICANT: KUBOTA, Michio
; APPLICANT: SUGIMOTO, Toshiyuki
; TITLE OF INVENTION: RECOMBINANT THERMOSTABLE ENZYME WHICH
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Browdy and Neimark
; STREET: 419 Seventh Street N.W. Ste. 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/505,377
; FILING DATE: 21-JUL-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 190180/1994
; FILING DATE: 21-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 109128/1995
; FILING DATE: 11-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP NOT YET RECEIVED
; FILING DATE: 04-JUL-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Browdy, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: MITSUZUMI-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-5197
; TELEFAX: (202) 737-3528
; TELEX: 249688
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 556 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-505-377-1

Query Match 58.7%; Score 1748.5; DB 2; Length 556;
Best Local Similarity 58.9%; Pred. No. 1.9e-143;
Matches 329; Conservative 87; Mismatches 136; Indels 7; Gaps 5;
QY 2 TFAYKIDGNEVFTLWAPYQSKVLKYLEKGLYEMERDEKGYFTITLNNVKVRDRYKYL 61
Db 3 SFGNIEKNKGIFKLWAPYVNSVKL-LSKKLIPMEKNDGFEVEIDDEENLTYSYII 61
QY 62 DDASEIDPPASRQPEGVHGPSQIIQESKEFNNETFLKKEDLIYIYHVTFTPEGTPEG 121
Db 62 EDREIDPPASRQPLGVHDKSQLIRTYQILDGLGVKVEDLIYIYHVTFTPEGTPEG 121
QY 122 VIRKLDYKLDGTTAIEIMPIAOPPGKRDGCVLYAVQNSYGGPGEKLVDEAHKK 181
Db 122 VIRKLDYKLDGTTGIELMPVAOPPGNRDNGYDGVLYAVQNTYGGPWEAKLVNEAHR 181
QY 182 GLGVILDVYVNHVGPENYMKLGPYSQYKTPWGLTFNFDADAESDEVKRFLENYEY 241
Db 182 GIAVILDVYVNHIGPEGNYLLGLGPYSDRYKTPWGLTFNFDGCGDQVRKFTLENYEY 241
QY 242 IKEYNVGDFRLDAVHAIDTSPKHILEEIAADVHKNYRIVIAESDLNDPRVNPKEKCY 301
Db 242 FKTFKIDGLRLDAVHAIFDNSPKHILQEIATKAHQLGVIAESDLNDPKIV--KDCGY 299
QY 302 NIDAQWDDFHHSHAYLTGERQYTTDFGNLDDIVKSKDVFYDGYKSNFRKTHGE 361
Db 300 KIDAQWDDFHHSHAVHAFITREKDYIYQDFGRIEDIEKTFKDVYDGYKSYRGRTHGAP 359
QY 362 VGLDGCNFFVYIQNDHQVGNRGKRIIKLVDRSYKIAAALYLLSPYIPMFMGEEY 421
Db 360 VGLDPPKRVFVYIQNDHQVGNRGKRIILTDKTTVMAATLVYLLSPYIPMFMGEEY 419
QY 422 BENPFFYFSDSKLIQVREGKRGKENGODTDPQDESTFNASKLSWKIDEEIFSFKIL 481
Db 420 ETNPFYFSDSKLIQVREGKRLKENQMDPQSEAFKLSKLSWKIDEEVLDYKQL 479
QY 482 IKMRKELSIACDRRVNVNGENWLIIGREYFSLYVFSKSSIEVKYSGTLLSSNNSFP 541
Db 480 INRKRYN-NCKRVKEVRREGNCITLIMEKIGIIASFDDIVINSKITGNLLIGI--GFPK 536
QY 542 HIEEGK-YEFDKGFALYKL 559
Db 537 KLKKDELKVNRGVGVYQL 555
RESULT 3
US-08-798-269-1
; Sequence 1, Application US/08798269
; Patent No. 6027918
; GENERAL INFORMATION:
; APPLICANT: MITSUZUMI, Hitoshi
; APPLICANT: KUBOTA, Michio
; APPLICANT: SUGIMOTO, Toshiyuki
; TITLE OF INVENTION: RECOMBINANT THERMOSTABLE ENZYME WHICH
; TITLE OF INVENTION: RELEASES TREHALOSE FROM NON-REDUCING SACCHARIDE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Browdy and Neimark
; STREET: 419 Seventh Street N.W. Ste. 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/798,269
; FILING DATE:

```

; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/505,377
; FILING DATE: 21-JUL-1995
; APPLICATION NUMBER: JP 190180/1994
; FILING DATE: 21-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 109128/1995
; FILING DATE: 11-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP NOT YET RECEIVED
; FILING DATE: 04-JUL-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Browdy, Roger L
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: MITSUZUMI-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-5197
; TELEFAX: (202) 737-3528
; INFORMATION FOR SEQ ID NO: 1:
; LENGTH: 556 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-798-269-1

Query Match      58.7%; Score 1748.5; DB 3; Length 556;
Best Local Similarity 58.9%; Pred. No. 1.9e-143;
Matches 329; Conservative 87; Mismatches 136; Indels 7; Gaps 5;

QY 2 TFAYKIDGNEVFTLWAPYQKSVKLVLEKGLYEMERDEKGYFTITLNNVKVRDRYKVL 61
DB 3 SFGNIENKKGIFKLWAPYVNSVKL-LSKKLIPMEKNDGEFFVEIDDEENLTSYII 61
QY 62 DDASEIPDPASRYOPEGVHGHSQIIOESKFENNETFLKEDLIIEHVGTFTPEGTFEG 121
DB 62 EDKREIPDPASRYOPLGVHDKSQIRTDYQILDGLGVKIEDLIIEHVGTFTPEGTFEG 121
QY 122 VIRKLDYLKDLGITAIEMPTAQPPGRDWDGYGVLYAVQNSYGGPEGFKLVDEAHKK 181
DB 122 VIEKLDYLKDLGITAIEMPTAQPPGRDWDGYGVLYAVQNTYGGPELAKLVNEAHR 181
QY 182 GLGVILDVNVNHVPEGNVMVKGYPFSQYKKTPTWGLTFFNPDDAESDEVRKFIENVEY 241
DB 182 GIAVILDVNVNHIGPEGNVLLGLPYFSDRYKTPWGLTFFNPDGRGCDQVRKFIENVEY 241

; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/505,377
; FILING DATE: 21-JUL-1995
; APPLICATION NUMBER: JP 190180/1994
; FILING DATE: 21-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 109128/1995
; FILING DATE: 11-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP NOT YET RECEIVED
; FILING DATE: 04-JUL-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Browdy, Roger L
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: MITSUZUMI-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-5197
; TELEFAX: (202) 737-3528
; INFORMATION FOR SEQ ID NO: 1:
; LENGTH: 556 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-055-210-1

Query Match      58.7%; Score 1748.5; DB 4; Length 556;
Best Local Similarity 58.9%; Pred. No. 1.9e-143;
Matches 329; Conservative 87; Mismatches 136; Indels 7; Gaps 5;

QY 2 TFAYKIDGNEVFTLWAPYQKSVKLVLEKGLYEMERDEKGYFTITLNNVKVRDRYKVL 61
DB 3 SFGNIENKKGIFKLWAPYVNSVKL-LSKKLIPMEKNDGEFFVEIDDEENLTSYII 61
QY 62 DDASEIPDPASRYOPEGVHGHSQIIOESKFENNETFLKEDLIIEHVGTFTPEGTFEG 121
DB 62 EDKREIPDPASRYOPLGVHDKSQIRTDYQILDGLGVKIEDLIIEHVGTFTPEGTFEG 121
QY 122 VIRKLDYLKDLGITAIEMPTAQPPGRDWDGYGVLYAVQNSYGGPEGFKLVDEAHKK 181
DB 122 VIEKLDYLKDLGITAIEMPTAQPPGRDWDGYGVLYAVQNTYGGPELAKLVNEAHR 181
QY 182 GLGVILDVNVNHVPEGNVMVKGYPFSQYKKTPTWGLTFFNPDDAESDEVRKFIENVEY 241
DB 182 GIAVILDVNVNHIGPEGNVLLGLPYFSDRYKTPWGLTFFNPDGRGCDQVRKFIENVEY 241
```

Qy	242	IKENVGDFRLDAVHAITDPSKPHILEETADVHKNRIVIAESDLNDPRVNPKECCY	301
Db	242	FKTKIDGLRLDAVHAIFDPSKPHILOETAERKAHQJGKVFIAESDLNDPKIIV--KDDCCY	299
Qy	302	NIDAQWVDDFHSHTHAYLTGEROQYTFDGNLDDIVKSKDVYDGYKSNFRRKTHGP	361
Db	300	KIDAQWVDDFHFAVHAIFTEKDYIYQDFGRIEDIEKTFDVFYDGYKSNFRGRTHGP	359
Qy	362	VGELDGCNFVYIIONHDQVGNRGKGERIKLVDRESYKIAAALYLSPIYPIFMFGEYGY	421
Db	360	VGDLPPRFVYFIQNHDDQVGNRNGENRSLITDKTYLMAATYILSPYIPLIFMGEYY	419
Qy	422	EENPFYFSDFSKLIQVREGRKXENGQDTPQDESTFNASKLSWKIDELIFEFYKIL	481
Db	420	ETNPFFFSDFSPVLKIGVREGRLKENQMDPQSEEAFLSKLSWKIDEEVLDDYKQL	479
Qy	482	IKNRKLSIACDRBVNVNGENLWIIKRGREFSYLVSFKSSTEVYSGTLLSSNNSFPQ	541
Db	480	INIRKRYN-NCKRKEVYREGNCITLIMEKIGIIASFDDIVNSKITGNLLTGI--GPFK	536
Qy	542	HIEGK-YEPDKGALYKIL	559
Db	537	KLAKKDELIKVNRGVGYQL	555

```

;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/607.321
; FILING DATE: 26-FEB-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/399,646
; FILING DATE: 07-MAR-1995
; APPLICATION NUMBER: JP 59834
; FILING DATE: 07-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 59840
; FILING DATE: 07-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: KUBOTA=5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 598 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-607-321-14

Query Match      35.6%; Score 1060.5; DB 1; Length 598;
Best Local Similarity 39.7%; Pred. No. 1e-83;
Matches 232; Conservative 92; Mismatches 197; Indels 63; Gaps 16;

Qy 1 MTFAYKIDGNEVI-----FTLWAPYQKSVKLKLEKG-LYEMER-----DEKGYFTI-- 46
Db 1 MHTYPREAAKPVILGPARYDVMAPNAESVTL--LAGGERYAMQRAETGPDAGWMTAAG 58

Qy 47 --TLNNVKVRDRYKYVLD-DASEIPDPASRYQPEGVHGPSQIIQESK-EFNNETFLKKE- 101
Db 59 APTDGNVD---YGYLLDGDETPLDPRTREQPDGVHVALSRFTDPSAYSWQDDAWQREL 114

Qy 102 -DLIIYEHVGTFTPEGTFFGVIRKLDYLKDLGTAIEIMPIAOPGKRDMGYDGVYLYA 160
Db 115 QGAVIYELHGLTFTPEGTLEAAAGKLDYLAGLVDFIELLPVNAFNGTHNGYDGVQWFA 174

Qy 161 VQSYGGPEGRKLVDEAHKKGVLVDVYVNHVGPENYMKLGPYFSQKYKTPWGLTF 220
Db 175 VHEYGGPEAFQFVDAHAAGLVQIDVYVNHGPSNYLPRFGPYLKQEGNTWGDV 234

Qy 221 NFDDAESDEVKRFILENVYVIKYNVDGFRDLDAVHALIDTSPKHILEEADVVVHKYNI 280
Db 235 NLDGPGSDHVRYYILDNLAMWLDRYVDGLDLDAVHALKDERAVHILEDGALADQISAE 294

Qy 281 V-----IAESDLNDPRVNPKEKCGYNIDQWDDFHHSIHAYLTGERQYTYDFGNLD 334
Db 295 VGRPLTLTAEADLNNRLLYPRDNGYGLGQWSDDFHHAHVNVNVTGTTGYSDFDLSA 354

Qy 335 DIVSKYDKVYVYDCKYSNFRKTHGEPVGLDGNF-----VYIQTNDQVGNRKG 386
Db 355 ALAKVLRDGFHDGYSYFSRHHGRPI-----NFSVHPAALVVCSDHQQIGNRATG 408

Qy 387 ERIKLVRESYKIAAALYLLSPYIMFIMGEYEEENPFYFSDFSKLIQVREGRK 446
Db 409 DRLSQTLPYGLAALAAVLTGPTPTMLLMGEYGAFTPWQFFTSHPPELKGKATAEGRI 468

Qy 447 KE---NGQD-----TDPQDESTFNASKWSKIDEE-----IFSFKYLIKMKKELSIAC-- 492
Db 469 KEFERMGWDPAVVPDPQDPETFRRSKLDWAEEAEGDHARLLELYRSLTALRRSTPDLTKL 528

Qy 493 ---DRRVNVNGENWLIKGREYFSLYVFSKSSIEVKYSGTLL 533
Db 529 GFEDTQVAFDEDAWRLRFRGGVQVLLNFSEQPVSLDAGAGTALL 572

;
; RESULT
; US-08-961-240-14
; Sequence 14: Application US/08961240
; Patent No. 5830715
; GENERAL INFORMATION:
; APPLICANT: KUBOTA, Michio
; APPLICANT: TSUSAKI, Keiji
; APPLICANT: HATTORI, Kazuko
; APPLICANT: SUGIMOTO, Toshiyuki
; TITLE OF INVENTION: DNA ENCODING ENZYME, RECOMBINANT DNA AND
; TITLE OF INVENTION: ENZYME, TRANSFORMANT, AND THEIR PREPARATIONS AND USE
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,240
; FILING DATE: 30-OCT-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/399,646
; FILING DATE: 07-MAR-1995
; APPLICATION NUMBER: JP 59834
; FILING DATE: 07-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 59840
; FILING DATE: 07-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: KUBOTA=5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 598 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-961-240-14

Query Match      35.6%; Score 1060.5; DB 2; Length 598;
Best Local Similarity 39.7%; Pred. No. 1e-83;
Matches 232; Conservative 92; Mismatches 197; Indels 63; Gaps 16;

Qy 1 MTFAYKIDGNEVI-----FTLWAPYQKSVKLKLEKG-LYEMER-----DEKGYFTI-- 46
Db 1 MHTYPREAAKPVILGPARYDVMAPNAESVTL--LAGGERYAMQRAETGPDAGWMTAAG 58

Qy 47 --TLNNVKVRDRYKYVLD-DASEIPDPASRYQPEGVHGPSQIIQESK-EFNNETFLKKE- 101
Db 59 APTDGNVD---YGYLLDGDETPLDPRTREQPDGVHVALSRFTDPSAYSWQDDAWQREL 114

Qy 102 -DLIIYEHVGTFTPEGTFFGVIRKLDYLKDLGTAIEIMPIAOPGKRDMGYDGVYLYA 160
Db 115 QGAVIYELHGLTFTPEGTLEAAAGKLDYLAGLVDFIELLPVNAFNGTHNGYDGVQWFA 174

Qy 161 VQSYGGPEGRKLVDEAHKKGVLVDVYVNHVGPENYMKLGPYFSQKYKTPWGLTF 220
Db 175 VHEYGGPEAFQFVDAHAAGLVQIDVYVNHGPSNYLPRFGPYLKQEGNTWGDV 234
```


OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/399,646
FILING DATE: 07-MAR-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 59834
FILING DATE: 07-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 59840
FILING DATE: 07-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: BROWDY, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: KUBOTA-5
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 597 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-399-646-4

Query Match 35.6%; Score 1059; DB 1; Length 597;
Best Local Similarity 40.5%; Pred. No. 1.4e-83;
Matches 229; Conservative 89; Mismatches 190; Indels 58; Gaps 15;

QY 14 FTLWAPYQKSVKLVLEKG-LYEMER-----DEKGYFTI-----TLNNVKVRDRYKYVLD 62
DB 18 YDWAAPNAESVTLL--LAGGERYAMQRAETGPDAGWMTAAGAPTDCGNVD-----YGYLLD 71
QY 63 -DASEIPDPASRYQPGVHGPSQIIQESK-EFNNETFLKKE--DLIIYETHVGTFTPEGT 118
DB 72 GDETPLPDPRTTRQPDGVHLSRTFDPSPAYSWQDDAQWQRELOGAVIYELHLGTFTEPGT 131
QY 119 FEGVIRKLDYKDLGTAIEMPIAOPGKRDMGYDGVLYAVONSYGGPEGRKLVDEA 178
DB 132 LEAAAGKLDYLAGLVDFIELLEVNAFNGTHNWGYDGVQWFAVHEAYGGPEAYQRFVDA 191
QY 179 HKKGLGVLDVYVNHVGPENYVVKLGPFYSQKYKTPWGLTFNFDAAESDEVRKFTILENV 238
DB 192 HAAGLVQIDVYVNHVGPENYVVKLGPFYSQKYKTPWGLTFNFDAAESDEVRKFTILENV 238
QY 239 EYWIKEYNVGDFRLDAVHAIDTSPKHILEEADVYVHKYNRIV-----IAESDLNDPRV 292
DB 252 AMWLDRYVDGLRLDAVHALKDERAVHILEDFCALADQISAEVGRPLTLIAESDLNPR 311
QY 293 VNPKEKCGYNIDAQWDDPHHSTHAYLTGRQGYTDFGNLDDIVKSYKDVYVYDQKYSN 352
DB 312 LYPRDNGYGLEQWSDDFHVAHVNVNVTGETTGYISDFDLSAALAKLVKRGDFHDSYSS 371
QY 353 FRKTHGPEVGLDGNF-----VYIQNHDQVGNKGGRIIKLVDRSYKTAAL 404
DB 372 FRERHGRPL-----NFSVHPAALVVCNSQNDHQIGNRATGDRLSQTLPGYSLAALV 425
QY 405 YLLSPYIPMIFMGEEYGEENPFYFFSDFSDSKLIQGVREGRKKE---NGQD-----TDPOD 457
DB 426 TLGTPPTMLLMGEEYGASTPWQFFTSHPPELGKATAEGRKEFERFMGWDPAVVPDQD 485
QY 458 ESTFNASKLSWKIDEE-----IPSFYKILMKRKEISAC-----DRRVNVNNGENWLI 507
DB 486 PEFRRSKLDWABAEGDHARLELYRSALTALRRSTPDLTKLGFEDTQVAFDEDAWLR 545
QY 508 KGREYFSLYVFSKSSIEVYSGTLL 533
DB 546 RRGGVQVLLNFSQPVSLDAGATALL 571

RESULT 10
US-08-607-321-4
Sequence 4, Application US/08607321
Patent No. 5716813
GENERAL INFORMATION:
APPLICANT: KUBOTA, Michio
APPLICANT: TSUSAKI, Keiji
APPLICANT: HATTORI, Kazuko
APPLICANT: SUGIMOTO, Toshiyuki
TITLE OF INVENTION: DNA ENCODING ENZYME, RECOMBINANT DNA AND
TITLE OF INVENTION: ENZYME, TRANSFORMANT, AND THEIR PREPARATIONS AND USE
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/607,321
FILING DATE: 26-FEB-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/399,646
FILING DATE: 07-MAR-1995
APPLICATION NUMBER: JP 59834
FILING DATE: 07-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 59840
FILING DATE: 07-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: BROWDY, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: KUBOTA-5
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 597 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-607-321-4

Query Match 35.6%; Score 1059; DB 1; Length 597;
Best Local Similarity 40.5%; Pred. No. 1.4e-83;
Matches 229; Conservative 89; Mismatches 190; Indels 58; Gaps 15;

QY 14 FTLWAPYQKSVKLVLEKG-LYEMER-----DEKGYFTI-----TLNNVKVRDRYKYVLD 62
DB 18 YDWAAPNAESVTLL--LAGGERYAMQRAETGPDAGWMTAAGAPTDCGNVD-----YGYLLD 71
QY 63 -DASEIPDPASRYQPGVHGPSQIIQESK-EFNNETFLKKE--DLIIYETHVGTFTPEGT 118
DB 72 GDETPLPDPRTTRQPDGVHLSRTFDPSPAYSWQDDAQWQRELOGAVIYELHLGTFTEPGT 131
QY 119 FEGVIRKLDYKDLGTAIEMPIAOPGKRDMGYDGVLYAVONSYGGPEGRKLVDEA 178
DB 132 LEAAAGKLDYLAGLVDFIELLEVNAFNGTHNWGYDGVQWFAVHEAYGGPEAYQRFVDA 191
QY 179 HKKGLGVLDVYVNHVGPENYVVKLGPFYSQKYKTPWGLTFNFDAAESDEVRKFTILENV 238
DB 192 HAAGLVQIDVYVNHVGPENYVVKLGPFYSQKYKTPWGLTFNFDAAESDEVRKFTILENV 238


```

Db      546  RRGVQVLLNFSEQVPSVLDGATALL 571
      RESULT 13
      US-08-399-646-2
      ; Sequence 2, Application US/08399646
      ; Patent No. 5536781
      ; GENERAL INFORMATION:
      ; APPLICANT: KUBOTA, Michio
      ; APPLICANT: TSUSAKI, Keiji
      ; APPLICANT: HATTORI, Kazuo
      ; APPLICANT: SIGIMOTO, Toshiyuki
      ; TITLE OF INVENTION: DNA ENCODING ENZYME
      ; TITLE OF INVENTION: ENZYME, TRANSFERASE
      ; NUMBER OF SEQUENCES: 18
      ; CORRESPONDENCE ADDRESS:
      ; ADDRESSES: BROWDY AND NEIMARK
      ; STREET: 419 Seventh Street, N.W.
      ; CITY: Washington
      ; STATE: D.C.
      ; COUNTRY: USA
      ; ZIP: 20004
      ; COMPUTER READABLE FORM:
      ; MEDIUM TYPE: FLOPPY disk
      ; COMPUTER: IBM PC compatible
      ; OPERATING SYSTEM: PC-DOS/MS-DOS
      ; SOFTWARE: PatentIn Release #1.0,
      ; CURRENT APPLICATION DATA:
      ; APPLICATION NUMBER: JUS/08/399 646

```

15;

```

; PRIOR REPLICATION DATA:
; APPLICATION NUMBER: JP 59840
; FILING DATE: 07-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: KUBOTA-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 589 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-399-646-2

Query Match 34.4%; Score 1025; DB 1; Length 589;
Best Local Similarity 40.3%; Pred. No. 1.2e-80;
Matches 229; Conservative 80; Mismatches 197; Indels 62; Gaps

QY 14 FTLWAPYQSKVLKVLKLGK-LYEMER-----DEKGYFTITLNNVKVRORYKVVLD-DAS 65
Db 10 FDIWAPGAVTUL--LGGERYEMRRPONGPADEGWTAADAPTADVDYGYLDDGDEI 67
QY 66 EIPDPASRYQPEGVHGPGSQIIQESKEFFNETFLKKEDL---IYEIHVGTFTPEGTFCV 122
Db 68 PLDPDTRRQPEGVHSLRTFDPGARHWQDAGWQRELQSGVIYELHIGTFTPEGLTDA 127
QY 123 IRKLDYLKDLGITAITMPTAOPPGKRDYGDGYLVYAVQNSYGGPEGPFKRLVDEAHKKG 182
Db 128 AGKLDYLAGLIDGFIELLPVNAFNGTHNNGYDGVQWFVAHEGYGGPAAVQRFVDAARAAG 187
QY 183 LGVILDVYVNHVGPENGYMKVGLGPFYSQYKTKPWGLTFTNFDDAESDEVKRFILENVYWI 242
Db 188 LGVIQDWWYNHLPSCNGLIPRYGPKLKHGEGNTWGPSVNLDPGSPDHVROYLTLDNVAMWL 247

```

16:

Db 536 RYWRGGVQVVLNFADRPISLDRPGTALL 563

Search completed: July 15, 2002, 12:18:04
Job time: 96 sec

[illegible]

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 12:16:28 ; Search time 32.43 Seconds
(without alignments)
1656.304 Million cell updates/sec

Title: US-09-298-924-6
Perfect score: 2978
Sequence: 1 MTFAYKIDGNEVIFTLWAPY.....PQHIEGKYEFDFKGFALYKL 559

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues
Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	2973	99.8	558	2 JC5135	alpha-amylase (EC
2	2467	82.8	561	2 S73087	alpha-amylase (EC
3	1748.5	58.7	556	2 JC5132	alpha-amylase (EC
4	1288.5	43.3	620	2 AH1827	alpha-amylase [imp
5	1060.5	35.6	598	2 S65770	maltooligosyltreha
6	1025	34.4	596	2 JC4697	trehalose trehaloh
7	1024.5	34.4	575	2 JC7727	malto-oligosyltreh
8	997.5	33.5	580	2 G70763	probable maltoolig
9	978.5	32.9	594	2 A10673	probable maltoolig
10	905.5	30.4	600	2 H75516	probable maltoolig
11	881.5	29.6	583	2 G83375	probable glycosyl
12	876.5	29.4	601	2 D96001	probable 1,4-alpha
13	860.5	28.9	595	2 AG3194	alpha-amylase Atu5
14	524.5	17.6	552	2 S19134	probable alpha-glu
15	524.5	17.6	552	2 AH1915	hypothetical prote
16	391	13.1	666	2 B56639	1,4-alpha-glucan b
17	374.5	12.6	627	1 S40048	1,4-alpha-glucan b
18	374.5	12.6	638	2 S18599	1,4-alpha-glucan b
19	365	12.3	1280	2 E95031	alkaline amylopull
20	364	12.2	1256	2 G97902	alpha-amylase (EC
21	361	12.1	630	2 D70363	1,4-alpha-glucan b
22	349.5	11.7	783	2 B84823	probable isoamylas
23	348.5	11.7	666	2 E71565	probable glyco-gen
24	347.5	11.7	764	2 AG1895	1,4-alpha-glucan b
25	345.5	11.6	422	2 S31839	1,4-alpha-glucan b
26	345.5	11.6	664	2 B86539	glycogen hydrolase
27	345.5	11.6	664	2 B72084	glycosyl hydrolase
28	345.5	11.6	666	2 G81717	glycosyl hydrolase
29	345.5	11.6	707	2 S77094	glycogen operon pr

30	344.5	11.6	741	1 S34218	1,4-alpha-glucan b
31	343	11.5	718	2 S73088	glycogen operon pr
32	340	11.4	730	1 I64118	1,4-alpha-glucan b
33	339.5	11.4	789	2 T01321	probable isoamylas
34	336	11.3	702	2 B85091	isoamylase-like pr
35	335	11.2	764	2 S47569	1,4-alpha-glucan b
36	335	11.2	1072	2 A84112	alkaline amylopull
37	334.5	11.2	717	2 H84057	pullulanase amyl l
38	333.5	11.2	720	2 B75540	glycogen operon pr
39	333.5	11.2	770	2 S76095	hypothetical prote
40	330.5	11.1	562	2 C82263	trehalose-6-phosph
41	329.5	11.1	818	2 T02231	probable isoamylas
42	329	11.0	733	2 T49679	probable branching
43	327.5	11.0	735	2 F98228	1,4-alpha-glucan b
44	327.5	11.0	735	2 AH3057	glycogen branching
45	326.5	11.0	603	2 AC0674	probable glycogen

ALIGNMENTS

RESULT 1

JC5135

alpha-amylase (EC 3.2.1.1) - Sulfolobus solfataricus

C:Species: Sulfolobus solfataricus

C:Date: 31-Jan-1997 #sequence_revision 27-Feb-1997 #text_change 20-Jun-2000

C:Accession: JC5135

R:Kobayashi, K.; Kato, M.; Miura, Y.; Kettoku, M.; Kameda, T.; Iwamatsu, A.

Biosci. Biotechnol. Biochem. 60, 1882-1885, 1996

A:Title: Gene cloning and expression of new trehalose-producing enzymes from the hype

A:Reference number: JC5134; MUID:97141610

A:Accession: JC5135

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-558 <KOB>

A:Cross-references: DBJ:D64130; NID:g987702; PIDN:BAAL1010.1; PID:g1838936

A:Experimental source: strain KM1

C:Comment: This enzyme hydrolyzes starch alpha-1,4 linkage adjacent to the alpha-1,1

C:Superfamily: trehalose trehalohydrolase

C:Keywords: glycosidase; hydrolase

Query Match 99.8%; Score 2973; DB 2; Length 558;

Best Local Similarity 100.0%; Pred. No. 1.1e-185;

Matches 558; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 TFAYKIDGNEVIFTLWAPYQSKVLKLVLEKGLYEMERDEKGYFTITLNNKVDRYKYVL 61

Db 1 TFAYKIDGNEVIFTLWAPYQSKVLKLVLEKGLYEMERDEKGYFTITLNNKVDRYKYVL 60

QY 62 DDASEIPDPASRYQPEGVHGSPQIIQESKEFNNEFLKKEDLIIYEHVGTFTPEGTFRG 121

Db 61 DDASEIPDPASRYQPEGVHGSPQIIQESKEFNNEFLKKEDLIIYEHVGTFTPEGTFRG 120

QY 122 VIRKLDYLKDLGITAITEIMPIAQFFPKRDMGVDGYLYAVQNSYSGSGEGFKRLVDEAHKK 181

Db 121 VIRKLDYLKDLGITAITEIMPIAQFFPKRDMGVDGYLYAVQNSYSGSGEGFKRLVDEAHKK 180

QY 182 GLGVILDDVYVNHVGPENGMVWKLGPYFSQKYTPMGLTFNFDADAESDEVRKFIENVEYW 241

Db 181 GLGVILDDVYVNHVGPENGMVWKLGPYFSQKYTPMGLTFNFDADAESDEVRKFIENVEYW 240

QY 242 IKEYNVGDFRLDAVHAIDTSPKHILEETADVVKHKNRIVIAESDLNDRVVPKCKCY 301

Db 241 IKEYNVGDFRLDAVHAIDTSPKHILEETADVVKHKNRIVIAESDLNDRVVPKCKCY 300

QY 302 NIDAQWDDFHSHIAYLTFGERQGYTDFGNLDDIVKSKYKDVFDGKYSNFRKTHGEP 361

Db 301 NIDAQWDDFHSHIAYLTFGERQGYTDFGNLDDIVKSKYKDVFDGKYSNFRKTHGEP 360

QY 362 VGEIDGCFVYVYIQNHQDVGNRGKGERIITKLVDRSEYKIAAALYLLSPYIPMFGEYVG 421

Db 361 VGEIDGCFVYVYIQNHQDVGNRGKGERIITKLVDRSEYKIAAALYLLSPYIPMFGEYVG 420

Qy	422	EENPFYFFSDFOSKLIQGVREGKKENGODTDPQDESTFNASKLSWKIDEEIFSFYKIL	481
Db	421	EENPFYFFSDFOSKLIQGVREGKKENGODTDPQDESTFNASKLSWKIDEEIFSFYKIL	480
Qy	482	IKMRKELSIACDRRVNVVNGENLLITKGREYFSLYVFSKSSIEVKYSGTLLSSNNSPQ	541
Db	481	IKMRKELSIACDRRVNVVNGENLLITKGREYFSLYVFSKSSIEVKYSGTLLSSNNSPQ	540
Qy	542	HIEEGKYEFDKGFALYKIL	559
Db	541	HIEEGKYEFDKGFALYKIL	558

```

RESULT      2
S73087      \
alpha-amylase (EC 3.2.1.1) precursor - Sulfolobus solfataricus
N:Alternate names: protein c0620
C:Species: Sulfolobus solfataricus
C:Date: 09-Oct-1997 #sequence_revision 24-Oct-1997 #text_change 20-Jun-2000
C:Accession: S73087
R:Sensen, C. W.; Klenk, H. P.; Singh, R. K.; Allard, G.; Chan, C. C. Y.; Liu, Q.
Mol. Microbiol. 22, 175-191, 1996
A:Title: Organizational characteristics and information content of an archaeal
A:Reference number: S73076; MUID:97055432
A:Accession: S73087
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-561 <SEN>
A:CROSS-references: EMBL:Y08256; NID:g1707679; PIDN:CAA69503.1; PID:g17076769
A:Experimental source: strain P2
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, Sep
C:Superfamily: trehalase trehalohydrolase
C:Keywords: glycosidase; hydrolase

```

DB 481 KSLIIKRIKEYGLACNRKLSVENGNGTWTYKGNCCLAAYVYFSKSVIEMKISGTLVLSSNS 540
 QY 539 FPOHIEEGKYEDKGFALYKL 559
 DB 541 FPSQITESKYELDKGFALYKL 561
 RESULT 3
 JC5132
 alpha-amylase (EC 3.2.1.1) - Sulfolobus acidocaldarius
 N:Alternate names: glycogenase,
 C:Species: Sulfolobus acidocaldarius
 C:Date: 31-Jan-1997 #sequence_revision 27-Feb-1997 #text_change 20-Jun-2000
 C:Accession: JC5132
 R: Kobayashi, K.; Kato, M.; Miura, Y.; Kettoku, M.; Komeda, T.; Iwamatsu, A.
 Biosci. Biotechnol. Biochem. 60, 1720-1723, 1996
 A:Title: Gene analysis of trehalose-producing enzymes from hyperthermophilic archaea
 A:Reference number: JC5131; MUID:97141330
 A:Accession: JC5132
 A:Status: preliminary; nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 1-556 <DOB>
 A:Cross-references: DDBJ:D64131; NID:q987704; PIDN:BAAL1011.1; PID:q987705
 A:Experimental source: strain ATCC33909
 C:Comment: This enzyme hydrolyzes glycosyltrehalose to liberate trehalose.
 C:Superfamily: trehalose trehalohydrolase
 C:Keywords: glycosidase; hydrolase
 Query Match 58.7%; Score 1748.5; DB 2; Length 556;
 Best Local Similarity 58.9%; Pred. No. 4.5e-106;
 Matches 329; Conservative 87; Mismatches 136; Indels 7; Gaps 5;
 QY 2 TFAYKIDNEVIFTWAPYQKSVKLKLEGLYEMERDEKGYFTTLNNVKVDRDKYKYL 61
 DB 3 SFGNGIEKNKGIFKLWAPYVNSVKLK-LSKKLIPMEKNDEGEFEVIDDIENLTYSI 61
 QY 62 DSAEIPDPASRYQEPYGVHPSQIIQESKEFNEFLFKEDLIIVIEIHVGTFTPEGTPE 121
 DB 62 EDKREIPDPASRYQPLGVHDKSQLRTDYIILDGLKVKIEDLIIVELHVGTSFQEGNFK 121
 QY 122 VIRKLDYKLDGITAIEIMPYIAQPGKRDWGDVYLVAVQNSYGGPEGFRKLVDIAHKK 181
 DB 122 VIEKLDYKLDGITGIELMPVAQPGNRDWDGVDGYLVAVQNTYGGPWELAKLVNEAHR 181
 QY 182 GLGVTLDDVYVNHVGPENVMYKIGPYFSQYKTPWGLTFNFDDESDERKVFLENVEY 241
 DB 182 GIAVTLDDVYVNHVGPENVLGLGYPFSRYKTPWGLTFNFDRCDDQVRKFIENVEY 241
 QY 242 IKEYNVGDGLDVAHAIIDTSPKHILEIADVYHKNYRIVIAESDLNDPRVYNPKECY 301
 DB 242 FKTFKIDGLRLDVAHAFDPSNPKILQETAEKAHQLGFVIAESDLNDPKIV--KDDCGY 299
 QY 302 NIDAQWDDFHSHAYLTGERQGYTDFGNLDDIVSKYKDVYVDGYSNPRKTHGEP 361
 DB 300 KIDAQWDDFHAVHAFITKEKDYIYDFGRIEDLEKTFKDYVDGYDSRYRGRTHGAP 359
 QY 362 VGELDGCNFVYVYIQNHQDVGNRKGERTIKVLDRSYIAAALYLLSPYIPMFGEY 421
 DB 360 VGDLPFRKFVVFVYIQNHQDVGNRNGERLSLTDKTYLMAATLYILSPYIPLFMGEY 419
 QY 422 EENPFYFSPSDSKLIGVREGRRKENGQDTFODESTFNASKLSWKIDEIFSFKYL 481
 DB 420 ETNPFYFSPSDPVLKGVREGRLKENQMDIPDSEAFKLSKLSWKIDEEVLDDYKOL 479
 QY 482 IKMRKELSIACDRRVYVNGENWLLIKGREYFSLVFSKSTEVYKSGTLLLSNSNP 541
 DB 480 INIKRYN-NCKRVKEVRREGNCITLIMEKIGIIASFDDIVINSKITGMLLGI--GPK 536
 QY 542 HIEEGK-YEFDKGFALYKL 559
 DB 537 KLKDKDELKVNKRGVGYOL 555

A: Molecule type: DNA
A: Residues: 1-620 <KUR>
A: Cross-references: GB:BA000019; PIDN:BAH77692.1; PTD:gl7135146; GSPDB:GN00179
A: Experimental source: strain PCC 7120
C: Genetics:
C: Gene: ali0168
C: Superfamily: trehalose trehalohydrolase

QY 47 - TLNNVVRDKRIKIVLD--DASLIPQFASRIQFVGHGFSQLIQESK -ERNNEITLRAE - 101

Dd 59 APTDGNVD---YGYLLDGETPLDPDTRRQPDGVHLSRTFDPAYSQWDDAQWGREL 114

115 QGAVIYELHGTPEPTLEAAGKLDYLAGLVDFIELLPVNAFNTHNWGYDGVQWFA 174

175 VHEAYGGPEAYQRFVDDAAHAAGLVQDVVYNHILGPSNGYLPFRGPYLKQSGENTWGDV 234
 Db

QY 221 NFDGSEVRKILLENVEIKLEINVDGKLDVAHALDLSFAHLLSELDVVRANKI 280
 | | | | | : | : | : | | | | : | | | : |
 | | | | | : | : | : | | | | : | | | : |
 Db 235 NLDGSGDHVRRYILDNLAMWLRYDVGDLRDVAHALKDERAVHILEDFGALQISAE 294

QY	V	-	AESDNP	KVNVN	KECUN	NIDAQW	DDFHHS	IHAITL	GREGQT	ITDRGNLD	334
		:	:	:	:	:	:	:	:	:	:
261	V	-	AESDNP	KVNVN	KECUN	NIDAQW	DDFHHS	IHAITL	GREGQT	ITDRGNLD	334
		:	:	:	:	:	:	:	:	:	:
		:	:	:	:	:	:	:	:	:	:
		:	:	:	:	:	:	:	:	:	:
Dbl	VGRPLTL	AESDLNNPRLLY	PRDVNGYGL	EQGWSDDFHHAVHVNTGETTGYYSDFSLA	354						

QY 335 DVASYKVDFVYDGKYNFRKTHGEPVGELDCNF-----VVYLONHDVGNRKGK 386
+ :
Db 355 ALAKVLRDGFFHDDGSYSFRRHHGRPI-----NFSAVHPAALVVCSONHHDIGNRATG 408

[illegible]

Oy	KE--NGQD---TDPODESTFNASKLSWIDEE-----IFSYKILIKMKRSELSIAC--	492
D6		
D7		
D8		
D9		
Dc		
Dd		
Df		
Dg		
Dh		
Db	KEFERMGWDPAVVDPDPQDPEIFRRRSKLWDAAEAEGDHARLLELYRSLTALRRSTPDITKL	528

Qy 493 -- -- DRRVVVNGENWLLIKGREYFSLYVFVSKSSIEVKVYSGTLLI 533
 : : : : : : : : : : : : : : : : : :
Db 529 GFEDTQVAFDEADRWLFRFGVGQVLLNFFSEQVPYSLDGAGTALL 572

	6
RESULT	
JC4 697	

C:Species: Rhizobium sp.
C:Date: 10-May-1996 #sequence_revision 19-Jul-1996 #text_change 20-Jun-2000
C:Accession: JC4697

Biosci. Biotechnol. Biochem. 60, 717-720, 1996
A: Title: Cloning and sequencing of trehalose biosynthesis genes from *Rhizobium* sp. M-
A: Reference number: JC4696; MUID:96210994

A;Molecule type: DNA

C:Superfamily: trehalose trehalohydrolase

Query Match 33.5%; Score 997.5; DB 2; Length 580;
Best Local Similarity 38.3%; Pred. No. 3.2e-57;
Matches 212; Conservative 91; Mismatches 215; Indels 35; Gaps 11;
QY 14 FTLWAPYQKSVKLVLEKGLYEMERDEKGYFTITLNNVKVRDRYKYVL-DDASEIPDPAS 72
DB 4 FRVWAPKALVRLDV-NGAVHAMRTSADGWHITV-AAPADARYGLLDDDPVLPDPRS 61
QY 73 RYQEGVHGSPQIIQESKEF---NNET---FLKEDLIIIEIHVGTTPGTEGVRKL 126
DB 62 AROPDGVHARSQRWEPQFGAARTDTGWPGRSVEGAVIYELHGTGTAGTDAATEKL 121
QY 127 DYLKDLGITAIEIMPIAQFPGKRDWGYDGYLVAVQNSYGGPFRKLVDKAEHKKGLGI 186
DB 122 DYLVDLGDIVELMPVNSFAGTGWGYDGYLVAVQNSYGGPFRKLVDKAEHKKGLGI 181
QY 187 LDVYVNHVGPENYMKLPGYFFSOKYKTPWGLTFNFDADAESDEVKRFILENVYWIKEYN 246
DB 182 IDAVFNHLGPGSNVLPREGPYLSSA-SNPWGDGINIAGADSDVHRHVIIDCALRWMDFH 240
QY 247 VDGFRLDAVHAIDTSPKHILEEIAADVHHYN-----RIVAESLNDPRVNVNPKKCG 300
DB 241 ADGLRLDAVHALVDTTAVHVLLEELANATRLSGQLRPLSLIAETDRNDPRLITRPSHG 300
QY 301 YNTDAQWVDFHSHIAYLGERGYTDFGNLDDIVKSYKDVGVYDGYKSNFRKTHGE 360
DB 301 YGTAQWDDIHHAIHTAVSGERGYYADFGSLATLAYLRNGYFHAGTYSFRRRRHGR 360
QY 361 PV--GELDGCNFVYIQNHQVGNRGKGERIIKLVDRSYKIAAALYLLSPYIPMFMGE 418
DB 361 ALOTSATPATRELLAYTCTHPQVGNRGLDRPSQYLTGGQAIKALTLGSPYTALEFMGE 420
QY 419 EYGEENPFYFFSDFSKLQGVREGKKGQD-----TDPQDESTFNASKLSWK-- 469
DB 421 EWGASSPFQFCRPELASHSTVAGRKEEFAEHGWAADIDPDPQPTQRCCKLNWAE 480
QY 470 ---IDEEIFSFYKLIKMRKELSTACD-----RNVNVNGENWLIKRGVYFSLYFSKS 521
DB 481 GSGEHARLHRFYLDLIALRNEADLPWLDHLMVDYDEQQRVWVRRGQLMNAE 540
QY 522 SIEVKSGTLLLS 534
DB 541 PTCVPVSGELVLA 553
RESULT 9
A:Title: probable hydrolase STY1503 [imported] - Salmonella enterica subsp. enterica serovar Typh
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: this species has also been called Salmonella typhi
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 27-Nov-2001
C:Accession: AI0673
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A:Reference number: AB0502; PMID:11677608
A:Accession: AI0673
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-594 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD01762.1; PID:g16052610; GSPDB:GN00176
C:Genetics:
A:Gene: STY1503
C:Superfamily: trehalose trehalohydrolase

Query Match 32.9%; Score 978.5; DB 2; Length 594;

Best Local Similarity 43.3%; Pred. No. 5.7e-56;
Matches 209; Conservative 76; Mismatches 169; Indels 29; Gaps 12;
QY 7 IDGNEVITLWAPYQKSVKLVLEKGLYEMERDEKGYFTITLNNVKVRDRYKYVLDDASE 66
DB 15 IADVVFRFLWATGQQQVLMRLAGKO-OEMQASGDGFTLDVSGVTPGTETYNFVLSDGMV 73
QY 67 IPDPASRYQEGVHGSPQIIQESKEFNNEFLAKK---EDLIIIEIHVGTTPGTEGFGVI 123
DB 74 LPDPASRAQKTDVNGPSYVIDPGSYAWRNTGWKSRWEQAVVYEMHTGTTPGTFHTAI 133
QY 124 RKLDYLKDLGITAIEIMPIAQFPGKRDWGYDGYLVAVQNSYGGPFRKLVDKAEHKKGL 183
DB 134 AKLPYLAELGVTVEVMPVAQFGRGWDGYDGYLVAVPHAYSAYGTPDDKAFIDTAHGYGL 193
QY 184 GVILDDVYVNHVGPENYMKLGP-YFSOKYKTPWGLTFNFDADAESDEVKRFILENVYWI 242
DB 194 SVILDIVLNHFGEENYLPPLAPAFHKERTPMGNGIAYD---VDARVYIIIEAPLYWL 250
QY 243 KEYNVDCGFRLDVAHAIIDTSPKHILEEIAADVHHK---YNRIVIAESLNDPRVNVNPKKC 299
DB 251 TEYHLDGLRDAIDQIEDSARHVLVEIAQRIREDITDRPIHLTTEDSRNIISLHPRDQD 310
QY 300 GYN--IDAQWVDFHSHIAYLGERGYTDFGNLDD--IVKSYKDVGVYDGYKSNFR 355
DB 311 GNAPLFTAENDDFHNNAVHVFATGETQAYYNDFADEPEKHLARALAEGFAYQGEIS---- 366
QY 356 KTHCEPVGELD---GCNFVYIQNHQVGNRGKGERIIKLVDRSYKIAAALYLLSPYI 411
DB 367 PQTEGPGVKGSTGPPVAFVDFIQNHQVGNRAQGLITLAGAERTKVLATLLLSPHI 426
QY 412 PMTFMGEEYGEENPFYFFSDFSKLQGVREGKKGK---NGQDT-DPQDESTFNASKL 466
DB 427 PLFMGEYGESRPFLETFDF-HCDLARAVREGRAKEFADHAGENVDPNAPETFORSKL 485
QY 467 SWK 469
DB 486 NWK 488
RESULT 10
H75516
A:Title: probable maltotoligosyltrehalose trehalohydrolase - Deinococcus radiodurans (strain R1
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
C:Accession: H75516
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896
A:Accession: H75516
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-600 <WHI>
A:Cross-references: GB:AE001905; GB:AE000513; MID:g6458144; PIDN:AAF10042.1; PID:g645
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR0464
A:Map position: 1
C:Superfamily: trehalose trehalohydrolase

Query Match 30.4%; Score 905.5; DB 2; Length 600;
Best Local Similarity 38.7%; Pred. No. 3.2e-51;
Matches 205; Conservative 86; Mismatches 170; Indels 69; Gaps 20;

QY 14 FTLWAPYQKSVKLVK-----LEKGLYEMERDEKGYFTITLNNVKVRDRYKYVLDDA 64
DB 38 FRLTWTARTAVRVNGTEHVTMTSLGGGIYELELP-----VGPGRALFVLDDGV 86
QY 65 SEIPDPASRYQEGVHGSPQIIQESKEFNNEF-----LKKEDLIIIEIHVGTFTP 115

```

Db      87 -PTDPVARFLPDGVGHGEAEVW---DFG--TFDWTDAWHGKTLADCVFEYVHVGTFTFP 139
Qy      116 EGTGEGVIRKLDYLKDLGITAIEIMTAQPPGKRDRGDCVLYAVQNSYGGPEGFRKLV 175
Db      140 EGTYRAAAEKLPLYLKGVTAIQVMPLEAFDQGGWGYDGAIFYAPYGRPEDLMAV 199
Qy      176 DEAHKKGVLGVILDVVYVNHVGPENYVWVKGAP-YFSQKYKTPNGLTFNFDASDEVKRFI 234
Db      200 DAAHRLGLGFLDVVYVNHFGPSNLYSSYAPSYFTDRFSAMGM--GLDYAE-PHMRVY 256
Qy      235 LENVEYTKENYVDGFRDLDAVHALIDTSPKHILEETADVHKY--NRIVIAESDLNDPRV 292
Db      257 TGNARMLRDLHFGLRLDLPATPYTDDSETHILTELAQEIHELGGTHLLLAEDHRLNLP 316
Qy      293 VNPKEKCYNIDAQWDDFHSHIHAYLTGEROGYYTDF-GNLDDIVKSYKDVFEVDGKYS 351
Db      317 VTVN----HLDGIWTDFFHETRVTLTGEGEYAGYRGGAELAYTIRGWRVYEGQFW 371
Qy      352 NFRKTH--GEPVELDGCNFFVYVYIQNHDOGVNKGRIIKL--VDRESYKIAAALYLL 407
Db      372 AVKGEERHGPSDALEAPNFVYCIQNHDDIGNRPLGERLHOSDGVTLHEYRGAALLI- 430
Qy      408 SPYTPMIFMGEEXEENPEVFESDFSCLKIQCVRGSKKE-----NGQDT-DPQDES 459
Db      431 -PWTPLLFQGOEAASTPFOFFSDHA-GEUQAVSGRKKFEGFGSGEDVDPQPAEQ 488
Qy      460 TFNASKLSWKIDE-----EIFSYKILIKMKRKLSTACDR-RVNVVNGEN 503
Db      489 TFLNSKLNWAEREGEHARTLRLYRLDLLRLREDPVLHNRQENLTGHD 538

RESULT 11
G83375
probable glycosyl hydrolase PA2164 [imported] - Pseudomonas aeruginosa (str
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: G83375
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hic
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; La
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an oppor
A:Reference number: A82950; MUID:20437337
A:Accession: G83375
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-583 <STO>
A:Cross-references: GB:AE004643; GB:AE004091; NID:g9948178; PIDN:AAG05552.1
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA2164
C:Superfamily: trehalose trehalohydrolase

```

Query Match	29.6%	Score 881.5;	DB 2;	Length 583;
Best Local Similarity	36.3%	Pred. No. 1.1e-49;		
Matches 203; Conservative	95;	Mismatches 174;	Indels 87;	Gaps 19;

Qy	3	FAYKIDGN - EVITFLWAPYQKSVKLVEKGLYEMERDEKGYFTITLNNVKVRDRYKYVL	61
		: : : : : : : : : : : : : : :	
Db	7	FGAQFQNGRTCFGLWAPDAREVRYETADGRDWPLEGSDSGWFATL - PCPPGTTRYRYR	65
		: :	
Qy	62	DDASEIPDPASRYQPEGVGPSQIIQES ----- KEFNNETFLKKEDLIYIEHVGTFTP	115
		: : : : : : : : : : : : : : : : : :	
Db	66	DGRPVPDPASQFPQPCGVHGHSQVLHDGTYANRWDEMRGPW -- HEAVIYELHVGLF --	120
		: :	
Qy	116	EGTTEGVIRKLDYLKDLGITAIMEPIAQPGPKRDWYDGVIYLVAVQNSYGSGPEGRKLIV	175
		: :	
Db	121	-GSVAEVERLPRLVELGYTAVELMPLGEFPGRRNWDGVLFPAPASAYGTPEQLKHIL	179
		: :	
Qy	176	DEAHKGLGVILDVYNHVGPEGINVMVK - LGPYFSQYKTKPTNGLTFNFDDAESDERKFII	234
		: : : : : : : : : : : : : : : :	

Db 180 DSAHGMGLMVFVDVIYNHFGDGNVLAQYAAAFRRDROTPWGQAIDF---RRGEVREFF 2330

Qy 235 LENVEYKIEYNVDGFRDLDAVHAIIIDTSPKHILEEIA-----DVVHKYNRIVIAE 284

Db 237 YENALMWLLDYRVDLDFDAVHAIPDSA--FLVEMARRLRGAAGPERHVH-----LVLE 288

Qy 285 SDLNDPRVVPKPKCGYNIDAQWDDPHHSIHAYLIGERQGYTDRGN-LDDIVKSYKDV 343

Db 289 NDDNRASLL---RQGY--DAQWDDGHGHALHVLLTGENDGYQDYPEPLRCLARLAE 342

Qy 344 FVYDGKYSNFRKTHGPEVGLDGCNFFVYIQNHDOVGNRGKGERIKLVLDRESYKIAAA 403

Db 343 FVYQGE-ANRHGRPRGEPESADLPADAFVLEQLQNHQDVGNRAFERLSVLAEPQALRLA 401

Qy 404 LYLSPYPIPMFGEYGEENPFYFSDFSKLIQVGERKKENGQD----- 452

Db 402 LQLLAPMIPILFEGEECAAREPFIYFTD-HQGLADAVREGRRKEGFGEFRFEGEGATLAS 460

Qy 453 -TDQDESTENASK-----LSWKIDDEEIFSFKYLIMKRWKELSI----- 490

Db 461 LPDPNAVETERSRGLADCDPAWR-----GFYRLLEIRHEHLPLRLRGARSLGVVTY 514

Qy 491 ---ACDRVNVVNGENWLI 506

Db 515 AGAALSARWLGDGSDWRI 533

RESULT 12

D96001

probable 1,4-alpha-glucan branching enzyme (EC 2.4.1.18) [imported] - Sinorhizobium meliloti

C:Species: Sinorhizobium meliloti

C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 14-Sep-2001

C:Accession: D96001

R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001

A:Title: The complete sequence of the 1.683-kb pSymB megaplasmid from the L. hymanii strain 239.

A:Reference number: A95842; MUID:21396508; PMID:11481431

A:Accession: D96001

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-601 <R>

A:Cross-references: GB:AL591985; PIDN:CAC49676.1; PID:gl5141163; GSPDB:GN000001

A:Experimental source: strain 1021, megaplasmid pSymB

R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Kopp, J.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; L.; Hyman, R.W.; Jones, T.

Science 293, 668-672, 2001

A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Kompheault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, A.; Title: The composite genome of the legume symbiont Sinorhizobium meliloti

A:Reference number: A96039; MUID:21368234; PMID:11474104

A:Contents: annotation

C:Genetics:

A:Gene: glgB2; Smb21447

A:Genome: plasmid

C:Superfamily: trehalose trehalohydrolase

C:Keywords: glycosyltransferase; hexosyltransferase

Query Match	29.4%	Score 876.5;	DB 2;	Length 601;
Best Local Similarity	35.3%	Pred. No. 2.5e-49;		
Matches 207; Conservative	99;	Mismatches 204;	Indels 77;	Gaps 19;

[illegible]

Db 137 ERLPHLAGAGITAIEIMPVQAQFPGVRGWGVDGLVHYAPHNAYGKPDLLKALVDAASHLGL 136
QY 184 GVLDDVYVNHVGPENVMYKLG-PYFSOKYKTPWGLTFNFDDAESDEVKRFILLENVEYWI 242
Db 197 TVLLDDVYVNHVGPENVMYKLG-PYFSOKYKTPWGLTFNFDDAESDEVKRFILLENVEYWI 253
QY 243 KEYNDVGFRLDAVHAII--DTSKPHILE---EADVVKYKRVIVIAESDLNDPRVVPKPK 299
Db 254 GHFRFGLRLDATEIQIRDTTKPHFLVALEHEVREAFARQIHLVLEDAHRRSLQORDAS 313
QY 300 GYNI--DAQWDDFHSHIAYLTGERQGYTDFGN--LDDIVKSYKDVYVYDGYKYSNFR- 354
Db 314 GARMLEDAWDDHNLHNVATGETKGYRLFADEPWGKIRSALEAGFAVPAKEDNFSP 373
QY 355 --RKTGCEPVGELDGCNFVYVYIQNDQVGNRGKGERIKLVDRSYKIAAALVLLSPYIP 412
Db 374 EGSRARVPQGRVN-----FLQNDQIGNRAFGERLASLQEDSLRVLAAMHMLTPQIP 427
QY 413 MIFMGEYGEENPFYFSDSFKLIQGVREGKKE-----NGQD-----TDQDEST 460
Db 428 LLFMGEYGETOPFYFSDY-QGEIAAAIRLGRRDEAENFGGLPECKTMDLLDPDLPDVP 486
QY 461 FNASKLSWK-----IDEEFYSFKILMKR----- 486
Db 487 FAGSKLRNRATSPAGERHLAVYRDLAVIRQHVPLIAGTAVPDGRAYETKDGIIAVDW 546
QY 487 ELSIAC-DRRVNVNNGENML-IKKGREYFSLVYFSKSSIEVKYSGTL 531
Db 547 QFGEACLELRVNLHETHAVPAIRGQPIFT---SETSSVETVYGSSEL 590

RESULT 13

AG3194
alpha-amylase Atu5284 [imported] - Agrobacterium tumefaciens (strain C58, Dupont) plasmid
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 01-Feb-2002
C:Accession: AG3194
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavlin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; PMID:11743193
A:Accession: AG3194
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-595 <KUR>
A:Cross-references: GB:AE008687; PIDN:AAL45973.1; PID:gl7743726; GSPDB:GN00188
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: Atu5284
A:Genome: plasmid
C:Superfamily: trehalose trehalohydrolase

Query Match 28.9%; Score 860.5; DB 2; Length 595;
Best Local Similarity 34.7%; Pred. No. 2.7e-48;
Matches 206; Conservative 89; Mismatches 246; Indels 53; Gaps 16;

QY 4 AYKIDGNEVIFTLWAPYQKSVKLVLEKGLYMERDEKGYFTITLNNVVRORYKYVLDD 63
Db 12 AQHQSGDFTDFINWAPASATVKL-WLNDAEFDHHTAGDGHIDIT-KPALPGDRYGVFLAD 69
QY 64 ASEIPDPASRYQPEGVHGPSQIIQESKEFNFTFLKK--EDLIYEHVGTFTPECTFEG 121
Db 70 RTRVADPASRQOEGRPGFSLVNHDFANKPNWKGPRWHEAVVYELHIGTFTPECTFAA 129
QY 122 VIRKLDYLDLGITALEIMPIAQFGPKRDWGDVGLYAVQNSYGGPEGRKLVDEAHKK 181
Db 130 AAKELEYLADVGTITIELMPLATFAGSRGWGDVGLQFSPQDYGTPDKELAFIDQAHGH 189

QY 182 GLGVILDDVYVNHVGPENVMYKLG-PYFSOKYKTPWGLTFNFDDAESDEVKRFILLENVEYWI 241
Db 190 GIMVLLDDVYVNHVGPAGNTLQAYAPAFKKHETPWGPAFDNFA---EYRSFLQNAFYW 246
QY 242 IKENYVDGFRDLDAVHAII--DTSKPHILE---EADVVKYKRVIVIAESDLNDPRVVPKPK 296
Db 247 LETYRFDGLRIDAADLAGDGEVDFLIEMAREVKRTIRNRHVHLVIEDARNAASPTM 306
QY 297 ERGCGYNIDQWDDFHSHIAYLTGERQGYTDFGN--LDDIVKSYKDVYVYDGYKYSNFR 354
Db 307 AGAILVDAQNDDFHVIHVATTNEEGGIYDFASRPYENLURRSIATGFGVVOGEPFRPR 366
QY 355 R-KTHGCEPVGELDGCNFVYVYIQNDQVGNRGKGERIKLVDRSYKIAAALVLLSPYIP 413
Db 367 NFAASGEPGSHLPFRFVFNFLHNDQAGNRLRGERLALIPPLFCTLEAILLCCPQTPL 426
QY 414 IFMGEYGEENPFYFSDSFKLIQGVREGKKE-----NGQD-----TDQDESTFN 462
Db 427 VFMGDEHGSANPFFPFDHPDHNRQEIRNRLKQAESFOGELPPDASQMVMDPNDQHTMQ 486
QY 463 ASKLSWKIDEEI-----FSFYKILMKR-----KELSIACDRRVNVNNGENW 504
Db 487 LSTLKWTHAETEGROSADMAALLAKRRGHVPLLCSHFEKGISLDCCEPRCLAI--DW 543
QY 505 LIKKGREYFSLVYFSKSSIEVKYSGTLSSNNNSFPQHIEEGKYE-FDKGFALY 557
Db 544 HFKAERLEMRANL-SENMCCELPAVKGELHRNGS-----VGNTRYEGYAAQFAIY 592

RESULT 14

S19134
probable alpha-glucanotransferase - Anabaena variabilis
C:Species: Anabaena variabilis
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 17-Mar-2000
C:Accession: S19134
R:Sato, N.
Plant Mol. Biol. 18, 165-170, 1992
A:Title: Cloning of a low-temperature-induced gene lti2 from the cyanobacterium Anaba
A:Reference number: S19133; MUID:92119230
A:Accession: S19134
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-552 <SAV>
A:Cross-references: EMBL:X59279; NID:g39252; PIDN:CAA41970.1; PID:g39253
C:Superfamily: neopolulanase, alpha-amylase core homology

Query Match 17.6%; Score 524.5; DB 2; Length 552;
Best Local Similarity 28.0%; Pred. No. 1.7e-26;
Matches 148; Conservative 90; Mismatches 162; Indels 129; Gaps 22;

QY 14 FTLWAPYQKSVKLVLEKGLYE--MERDEKGYFTITLNNVVRD---RYKYVLDDAS--- 65
Db 7 FTLFAPYKGAALIASFSQWQIMPKKGGDGYFRTT---VELEDGTQYKFRVQTRSWFF 63
QY 66 -----EIPDPASRYQPEGVHGPSQIIQESKEFN--ETFLKKED-----LIYET 108
Db 64 EEDQWVDVTDPTATDIDES-SKDNSIARIKDKGEKIVYVWQHDQKPLPADHELVIYEL 122
QY 109 HVGTFE-----PECTFEGVIRKLDYLDLGITALEIMPIAQFGPKRDWGDVGLYAV 161
Db 123 HVGDFSGGDDPYARGKYKHVIEKLDYLCLELGINALELLPVKEYPGDYSWGNPRYFFAT 182
QY 162 QNSYGGPEGRKLVDEAHKKGLGVILDDVYVNHVGPENY-MVKLGYFSQKYKTP---WG 217
Db 183 ESSYGSTADLKLVLDCHQGRIRIIMDGIYNHSEASSPLTQIDHDYVHHEPRDPDNNWG 242
QY 218 LTFNF---DDAESDEVKRFILLENVEYWKYKYNVDGFRDLDAVHAII--DTSKPHILE---BE 269
Db 243 PFNVEYHEDNLETYPARKFIDGTVRYWVGVEYHLDGIRYDAARQIANYDFMHWIAQEA 302
QY 270 IADVVHKYNRIVIAESDLNDPRVNV---PKKCGYVNIADQWVDDFHHSTHAYLTGERQGY 326

Db 303 TAGAKPFYN---VAEHIPETTSITNLDGPMDCG-----WHDSPYHTIKAHICGDT--- 349
Qy 327 YTFGNLDDIVKSKYKDVYVYDGKYSNFRKTHGEPVGELDCGNFVYIQNHQ-----QV 380
Db 350 -FDLENLKDVI-----DPKRGQF-----LGATNVVNYLTNHDHIMVEL 388
Qy 381 GNRGGERIILKVDRESY---KIAAALYLLSPYIPMIFMGEEYGEENPFYFSDSKL 437
Db 389 GNR-----EIFHDEAFRAKLGTAIIMTAVGVPLIWMGEFGEYKP----- 429
Qy 438 IQGVREGRRKKGQDTPDQDESTFNASKLSWKIDEEIFSFYKILIKMRK 486
Db 430 -----KQDQSKIDWTLLGNDLNRSLFDYHKGLIGLRK 462

RESULT 15

AH1915
hypothetical protein all10875 [imported] - Anabaena sp. (strain PCC 7120)
C:Species: Anabaena sp.
A:Note: Anabaena sp. (strain PCC 7120) is a synonym of Nostoc sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 11-Jan-2002
C:Accession: AH1915
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriquchi
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AH1915
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-552 <KUR>
A:Cross-references: GB:BA000019; PIDN:BAE72832.1; PID:gl7130220; GSPDB:GN00179
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: all10875
C:Superfamily: neopullulanase; alpha-amylase core homology

Query Match 17.6%; Score 524.5; DB 2; Length 552;
Best Local Similarity 28.0%; Pred. No. 1.7e-26;
Matches 148; Conservative 90; Mismatches 162; Indels 129; Gaps 22;

Qy 14 FTLNAPYQKSVKLVLEKGLYE--WERDEKGYFTITLNNVKVRD---RYKVLDDAS--- 65
Db 7 FTLFAPYKGAALIASFSDMQEIPMKGGDGYFRT---VELEDGYQYKFRVQTRSWFF 63
Qy 66 -----EIPDPASRYQPEGVHGPSQIIQESKEFNN--ETFLKKED-----LIIVEI 108
Db 64 EEDQWVDVTDYATIDIDES--SGKNSIARIKDGKIVDTYVQWQHDKPLPADHELVIYEL 122
Qy 109 HVGTFT-----PEGTGFEVIRKLDYKDLGITAIEIMPIAQPPGKRDWGDGVLYAV 161
Db 123 HVGDFSGGEDDPYARGVKYKVKIEKLDYLCELGINAIELLPVKYKPGDYSWGNPRYFFAT 182
Qy 162 QNSYGGPEGFRKLVDEAHKGLGVILDVVNNHVGPEGNY--MVKLGPYFSOKYKTP---WG 217
Db 183 ESSYSGSTADLKKLVDECHQRGIRIIMDGIYNHSEASSPLAQIDHDYWHHEPRDPDNNWG 242
Qy 218 LTFNF---DDAESDEVKFTLENVEYKWEYNVVDGFRDLDAVHAIIDTSPKHIL---EE 269
Db 243 PEFNVEHYDENLEYTPARKFIGDVTYVWGEYHLDGIRYDAAQIAIYDFMHWIAQAKK 302
Qy 270 IADVVKYINRIVIAESDLNPRVVN---PKEKGCYNIDAQWDDFHHSIHAYLTGERQGY 326
Db 303 TAGAKPFYN---VAEHIPETTSITNLDGPMDCG-----WHDSPYHTIKAHICGDT--- 349
Qy 327 YTFGNLDDIVKSKYKDVYVYDGKYSNFRKTHGEPVGELDCGNFVYIQNHQ-----QV 380
Db 350 -FDLENLKDVI-----DPKRGQF-----LGATNVVNYLTNHDHIMVEL 388
Qy 381 GNRGGERIILKVDRESY---KIAAALYLLSPYIPMIFMGEEYGEENPFYFSDSKL 437
Db 389 GNR-----EIFHDEAFRAKLGTAIIMTAVGVPLIWMGEFGEYKP----- 429

Qy 438 IQGVREGRRKKGQDTPDQDESTFNASKLSWKIDEEIFSFYKILIKMRK 486
Db 430 -----KQDQSKIDWTLLGNDLNRSLFDYHKGLIGLRK 462

Search completed: July 15, 2002, 12:18:43
Job time: 135 sec

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	1060.5	35.6	598	1	TREZ_ARTSP	Q44316 arthrobacte
2	1025	34.4	596	1	TREZ_RHISP	Q53238 rhizobium s
3	1024.5	34.4	575	1	TREZ_ARTRM	Q9a9n6 arthrobacte
4	997.5	33.5	580	1	TREZ_MYCTU	Q10769 mycobacteri
5	965	32.4	589	1	TREZ_BREHE	Q52520 brevibacter
6	391.5	13.1	666	1	GLGB_BACCL	P30537 bacillus ca
7	374.5	12.6	627	1	GLGB_BACSU	P39118 bacillus su
8	374.5	12.6	639	1	GLGB_BACST	P30538 bacillus st
9	340	11.4	730	1	GLGB_HAEIN	P45177 haemophilus
10	335	11.2	764	1	GLGB_STRAU	P52980 streptomyce
11	333.5	11.2	770	1	GLGB_SYNY3	P52981 synechocyst
12	326.5	11.0	773	1	GLGB_SYNP7	P16954 synechococc
13	326	10.9	731	1	GLGB_MYCTU	Q10625 mycobacteri
14	321	10.8	639	1	GLGB_PUTFI	P30539 butyrivibri
15	318	10.7	659	1	GLGX_HAEIN	P45178 haemophilus
16	318	10.7	734	1	GLGB_AGRU	P52979 agrobacteri
17	314	10.5	843	1	PUIA_THEMEA	O33840 thermotoga
18	303	10.2	561	1	TREZ_BACSU	P39795 bacillus su
19	300	10.1	498	1	AMY3_DICTH	P14899 dictyoglo
20	300	10.1	581	1	MAX5_YEAST	P40884 saccharomyc
21	299	10.0	520	1	AMY_BACME	P20845 bacillus me
22	298.5	10.0	562	1	O166_BACTR	P29094 bacillus th
23	298	10.0	558	1	O166_BACCE	P21332 bacillus ce
24	298	10.0	562	1	AMY2_DICTH	P14898 dictyoglo
25	294.5	9.9	574	1	CDAS_THEET	P29964 thermoanaer
26	294.5	9.9	728	1	GLGB_ECOLI	P07762 escherichia
27	288.5	9.7	861	1	GLGB_SOLU	P30924 solanum tub
28	287	9.6	657	1	GLGX_ECOLI	P15067 escherichia
29	285.5	9.6	702	1	GLGB_HUMAN	Q04446 homo sapien
30	283	9.5	963	1	TREZ_THEAQ	O06458 thermus aqu
31	282.5	9.5	555	1	O166_BACCO	Q45101 bacillus co
32	282.5	9.5	799	1	GLGB_MAZE	Q08047 maza maye
33	278	9.3	820	1	GLGB_ORYSA	Q01401 oryza sativ

Db 59 APTDGNVD---YGLLDGDETPDPTRRQPDGCVHALSRFTDPSAYSQWDDAQWQREL 114
Qy 102 -DLIIIEHVGFTEPGEFVIRKLDYLDKLGITATIEPIAOPPGKRDWGDYGVLYA 160
Db 115 QGAVIYELHGTFTPEGTELEAAAGKLDYLAGLVDFIELLPVNAFNGTHNGYDGVQWFA 174
Qy 161 VQNSYGGPGEFRKLVDEAHKGLGVLVDVYVNHVGPENVMVKGYPFSQKYKTPWGLTF 220
Db 175 VHEAYGPEAYQRFVDAHAAGLVQDVVYVNHVGPENVMVKGYPFSQKYKTPWGLTF 234
Qy 221 NFDAESDEVKFFILENVYVYIKYNDVDFRLDAVHALIDTSPKHLIEEADVYVHKYNI 280
Db 235 NLDGPGSDHVRVYILDNLAMWLRDYRVDDLDAVHALKDERAVHILEDFGALADQISAE 294
Qy 281 V-----IAESDLNDPRVVPKCEKGVNIDAQWDDFHHSIYALITGERQGYTDFGNLD 334
Db 295 VGRPLTLIAESDLNPNRLLYPRDVNGVLEQWSDDFHHAHVNVVGTGTYTDFDLSA 354
Qy 335 DIVSKYKDVYVYDGYKSNFRKTHGEPVGLDGCNF-----VYIQNHQDVGNRKGK 386
Db 355 ALAKVLRDGFPHDGSYSFSRHHGRPI-----NFSAVHPAALVVCSSQNHQDIGNRATG 408
Qy 387 ERIIKLVDRYSKYAAALYLLSPYIPMFWMGEYGEENPFYFSDFSKLIQVREGK 446
Db 409 DRLSOTLPYGLSALAVALTLTGPTPMLLMGEYGAFTPWQFFTSHPPELKGATAEGRI 468
Qy 447 KE-----TDPODESTFNASKLSWKIDEE-----IPSFYKILIKMKKELSIAC-- 492
Db 469 KEFERMGWDPAVVPDQDPQETFRSKLDWAEAGDHARLLELYRSITALRRSTPDITKL 528
Qy 493 ---DRRVNVNGENWLIKIGREYFSLYVFSKSSIEVYSGTLL 533
Db 529 GFEDTQVAFEDARWLFRFRGGVQVLLNFSQVPSYLDGAGTALL 572

RESULT 2

ID TREZ_RHISP STANDARD; PRT; 596 AA.
AC Q53238;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Malto-oligosyl trehalose trehalohydrolase (EC 3.2.1.141) (MTHase) (4-
alpha-D-[(1->4)-alpha-D-glucano]trehalose trehalohydrolase)
DE (Maltooligosyl trehalose trehalohydrolase).
GN TREZ.
OS Rhizobium sp. (strain M-11).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=391;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96219094; PubMed=8829547;
RA Maruta K., Hattori K., Nakada T., Kubota M., Sugimoto T., Kurimoto M.;
RT "Cloning and sequencing of trehalose biosynthesis genes from Rhizobium
sp. M-11.";
RL Biosci. Biotechnol. Biochem. 60:717-720(1996).
CC -1- CATALYTIC ACTIVITY: Hydrolysis of alpha-(1->4)-D-glucosidic
linkage in 4-alpha-D-[(1->4)-alpha-D-glucanosyl](n) trehalose to
CC yield trehalose and alpha-(1->4)-D-glucan.
CC -1- PATHWAY: Trehalose biosynthesis.
CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>;
CC or send an email to license@isb-sib.ch).
CC

DR EMBL; D78001; BA01187.1; -.
DR InterPro; IPR000461; Alpha-amylase.
DR Pfam; PF00128; alpha-amylase; 1.
KW Hydrolase; Glycosidase.
FT ACT_SITE 265 265 BY SIMILARITY.
SQ SEQUENCE 596 AA; 65262 MW; 73EA90AE0534DDCD CRC64;

Query Match 34.4%; Score 1025; DB 1; Length 596;

Best Local Similarity 40.3%; Pred. No. 8.4e-60;
Matches 223; Conservative 197; Indels 62; Gaps 16;

Qy 14 FTLWAPYQKSYKLVLEKG-LYEMER-----DEKGYFTITLNNVYKVRDRYKYVLD-DAS 65
Db 17 FDIWAPAGTTL--LAGGERYEMRRRPGNGPADGWTAAADAPTGDADVDYGYLLDGEI 74
Qy 66 EIPPASRYQPEGVHGSQIIQESKEFNNEFLKEDL---IYIEIHVGFTEPGEV 122
Db 75 PLDPRTTRRQPEGVHLSRTFDPGAHRWQDAGWQRELGQSVIYELHGTFTPEGLDAA 134
Qy 123 IRKLDYLDKLGITATIEPIAOPPGKRDWGDYGVLYAVQNSYGGPGRKLVDEAHKKG 182
Db 135 AGKLDYLAGLIDFIELLPVNAFNGTHNGYDGVQWFAVHGEYGGPAAQRFVDAHAAG 194
Qy 183 LGVILDVYVNHVGPENVMVKGYPFSQKYKTPMGLTFNFDAAESDEVKFFILENVYI 242
Db 195 LGVIQDVVYVNHVGPENVMVKGYPFSQKYKTPMGLTFNFDAAESDEVKFFILENVYI 254
Qy 243 KEYNVDFRLDAVHALIDTSPKHILEE---IADVVHKYN---RIVIAESDLNDPRVVPK 296
Db 255 RYRVDFGLRLDAVHALKDERAVHILEEFGALADALSSEGGRLTLIAESDLNPNRLLYPR 314
Qy 297 EKCYNIDAAQWDDFHHSIYALITGERQGYTDFGNLDIVKSKYDVYVYDGYKSNFRK 356
Db 315 DVNGYGLAGQWSDDFHHAHVNVVSGETGYSDFSGLAKALVLDGFFHDSYSSRGR 374
Qy 357 THGEPVGLDGCNF-----VYIQNHQDVGNRKGRIIKLVDRYSKYIAAALYLLS 408
Db 375 CHGRPI-----NFSAVHPAALVVCSSQNHQDIGNRATGDRLSQSLPYGLSALAVALTLTG 428
Qy 409 PYIPMFWMGEYGEENPFYFSDFSKLIQVREGKKE-----NGQD-----TDPODESTF 461
Db 429 PFTPLMFWMGEYGAFTPWQFFTSHPPELKGATAEGRIREFRMGWDPAVVPDQDPETF 488
Qy 462 NASKLSK-----IDEEIFGFYKILIKMKR-----ELSIACDRRVNVNGENWLIKGR 510
Db 489 TRSKLDWAEASAGDHARLLELYRSITLRLRSTPELARLGFA-DTAVEFDDARWL----- 542
Qy 511 EYFS-----LYVFSKSSIEVYSGTLL 533
Db 543 RYWRGGVQVVLNFAADRPISLDRPGTALL 570

RESULT 3

ID TREZ_ARTRM STANDARD; PRT; 575 AA.
AC Q9AJN6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Malto-oligosyl trehalose trehalohydrolase (EC 3.2.1.141) (MTHase) (4-
alpha-D-[(1->4)-alpha-D-glucano]trehalose trehalohydrolase)
DE (Maltooligosyl trehalose trehalohydrolase).
GN TREZ.
OS Arthrobacter ramosus.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Micrococciaceae; Micrococcaceae; Arthrobacter.
OX NCBI_TaxID=1672;
RN [1]
RP SEQUENCE FROM N.A.
RC SPRAIN=S34;
RA Yamamoto T., Maruta K., Watanabe H., Yamashita H., Kubota M.,
RA Fukuda S., Kurimoto M.;

```
RT "Trehalose producing operon treyZ from Arthrobacter ramosus S34.";  
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
CC -!- CATALYTIC ACTIVITY: Hydrolysis of alpha-(1->4)-D-glucosidic  
CC linkage in 4-alpha-D-[(1->4)-alpha-D-glucanosyl](n) trehalose to  
CC yield trehalose and alpha-(1->4)-D-glucan.  
CC -!- PATHWAY: trehalose biosynthesis.  
CC -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO  
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC -----  
DR EMBL; AB045141; BAB40766.1; -;  
DR InterPro; IPR000461; Alpha_amylase.  
DR Pfam; PF00128; alpha-amylase; 1.  
KW Hydrolase; Glycosidase.  
FT ACT_SITE 250 250 BY SIMILARITY.  
SQ SEQUENCE 575 AA; 63079 MW; 284A3F20207E228B CRC64;  
  
Query Match 34.4%; Score 1024.5; DB 1; Length 575;  
Best Local Similarity 39.4%; Pred. No. 8.6e-60;  
Matches 215; Conservative 90; Mismatches 189; Indels 51; Gaps 13;  
  
QY 14 FTLWAPYQSVKLVLEKGLYE--MERDEKGYFTIFL---NNVYRDRYKYYLDDASEIP 68  
Db 5 FFWAPAAQVTL-VVQGGAELPLTRDENGWALQPPWDGGPDLVD-YGYLVDGKGPEA 62  
QY 69 DPASRYQPEGVCHGPSIIQESKEFNTEFLKED-----LIYEIHVGTFTPEGTF 119  
Db 63 DPLSLRQPPGVH-----ELGREFDPARYAGDGGWRGLDGAIVYELHVGTFTPEGTL 116  
QY 120 EGYIRKLDYKIDGITAETIMPTAQPPGKRDWGYDGVLYAVONSYGGPGFKRLKDEAH 179  
Db 117 DSAIRDLHLVRLGVDVAVELLPVNAFNGHWGVDGLVAVHVPYGGPEAQRFVDACH 176  
QY 180 KGLGVILDVYNNHVGPEGNVYKLGYPFSQYKKTWPGTLTFNPDAAESDEVKRFLENYE 239  
Db 177 ARG LAVVQDVYNNHVGPEGNVYKLGYPFSQYKKTWPGTLTFNPDAAESDEVKRFLENYE 236  
QY 240 YMIKEYNVDFRLDVAHVIDTSPKHILEIA----DVVHKYR--IVTAESDLNDPRVY 293  
Db 237 YLDRMHADGLRLDVAHLDARALHLEELAAARVDDELAGELGRLPTLIAESDLNDPKLI 296  
QY 294 NPKEKGYNIDAQWVDDFHSHAYLTGRQGYTDFGNLDDITVSKDVFVYDGYKSNF 353  
Db 297 RSRAAHGYGLDAQWDDVHVAHVHANYGTGVGYADFGGLGALVKYVQRFQWFDGTWSF 356  
QY 354 RKKTGEPVG-ELDGCNFVYVYQNDQVGNRKGKRIKLVDRSYKIAAALYLSPTYIP 412  
Db 357 RERHHGRPLDDIPFRLFAFAQHDQVGNRAVGDMSAQVGGEGSLAAALVLLGPFTP 416  
QY 413 MLFPMGEYGEENPFYFSDSKLQGVREGKKE---NGQD----TPQDESTFNASK 465  
Db 417 MLFPMGEWGARTPQWQFTSHPELGEATARGIAEFARMGWDPAVVPDPQATFARSH 476  
QY 466 LSWKIDEE-----IFSFKYILMKRKLSTIACDRNNVYNGENWLLIKGREYFSLVFSK 520  
Db 477 LQNSEPERPHAGLAFYDTLALRELVPDAPARE-----VDADAEARGVFAFSR 526  
QY 521 SSIEV 525  
Db 527 GPLRV 531  
  
RESULT 4  
TREZ_MYCTU STANDARD; PRT; 580 AA.
```

```
AC Q10769;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Malto-oligosyltrehalose trehalohydrolase (EC 3.2.1.141) (MTHase) (4-  
DE alpha-D-[(1->4)-alpha-D-glucano]trehalose trehalohydrolase)  
DE (Maltooligosyl trehalose trehalohydrolase).  
GN TREZ OR RV1562C OR MT1613 OR MTCY48.03.  
OS Mycobacterium tuberculosis.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.  
OX NCBI_TaxID=1773;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=H37RV;  
RX MEDLINE=98295987; PubMed=9634230;  
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
RA Gordon S.V., Eiglmeyer K., Gas S., Barry C.E. III, Tekala F.,  
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,  
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,  
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,  
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;  
RA "Deciphering the biology of Mycobacterium tuberculosis from the  
RT complete genome sequence.";  
RL Nature 393:537-544(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CDC 1551 / Oshkosh;  
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,  
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,  
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,  
RA Delcher A., Uterback T., Weidman J., Khouri H., Gill J., Mikula A.,  
RA Bishai W.;  
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and  
RT laboratory strains.";  
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
CC -!- CATALYTIC ACTIVITY: Hydrolysis of alpha-(1->4)-D-glucosidic  
CC linkage in 4-alpha-D-[(1->4)-alpha-D-glucanosyl](n) trehalose to  
CC yield trehalose and alpha-(1->4)-D-glucan.  
CC -!- PATHWAY: Trehalose biosynthesis.  
CC -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO  
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC -----  
DR EMBL; Z74020; CAA989329.1; -;  
DR EMBL; AE007027; AAK45880.1; ALT_INIT.  
DR TIGR; MT1613; -;  
DR Tuberculist; Rv1562c; -;  
DR InterPro; IPR000461; Alpha_amylase.  
DR InterPro; IPR004193; isoamylase_N.  
DR Pfam; PF00128; alpha-amylase; 1.  
DR Pfam; PF02922; isoamylase_N; 1.  
KW Hydrolase; Glycosidase; Complete proteome.  
FT ACT_SITE 247 247 BY SIMILARITY.  
SQ SEQUENCE 580 AA; 64076 MW; 506190468F44B862 CRC64;  
  
Query Match 33.5%; Score 997.5; DB 1; Length 580;  
Best Local Similarity 38.3%; Pred. No. 5.1e-58;  
Matches 212; Conservative 91; Mismatches 215; Indels 35; Gaps 11;  
  
QY 14 FTLWAPYQSVKLVLEKGLYEMERDEKGYFTIFLNNVYRDRYKYYL-DDASEIPDPAS 72  
Db 4 FRWAPKPALVRLDV-NGAVHAMTSADGWHTTV-AAPADARYGYLLDDDDPTVLDPKRS 61
```


Query Match	12.6%;	Score 374.5;	DB 1;	Length 627;
Best Local Similarity	23.1%;	Pred. No. 3e-17;		
Matches 139;	Conservative 94;	Mismatches 223;	Indels 147;	Gaps 24;
QY	14	FTLWAPYQKSVKLKVLKGLYE-----MER-DEKGYFTITLNNVVRDXY-VLDD	63	
DB	41	FCWVAPHASERV-----AGDFNSGEEHVRVNDNGITLFPIGICEKERYEIVTN	96	
QY	64	ASBI-----PDPA-S-RYQEPGHSPQIIQESKEFNNEFLFKEDLIIVEI	108	
DB	97	NGEIRLKADPYYATYSEVRNTASLTVDLEGYSWQDKWQKQAKT---LYEKPFFVIEL	153	
QY	109	HVCTFTEGTFEGVIRK-----LDYKDLGTAITEIMPIAQFPGKRDWGDGVLYAVQ	162	
DB	154	HLSGSKKSHDGRHSYKELSQTLIPYIKKHGTHTELLPVEYHVDYRSGYQGTGYISPT	213	
QY	163	NSYGGPEGFRKLVDAAHKKGLGVLDVVVNVH--GPEGNMVMLGCPYFSQKYKTP-----	215	
DB	214	SREGPHDLMKFVDECHOQIGVILDWVPGHCKDAHGLYMFDEGPELYEYKEERDRENWL	273	
QY	216	WGTLTFNFDAAEDEVKRFLENYVEWIKYENVVDGFRDLDAVHAI-----DTSKPHILEEAD	272	
DB	274	WG-TANF-DLGKPEVHSFLISNALYWAEPYHIDGFRDVAVANILYWPNOQDERHTNPYAVD	331	
QY	273	VVHKYNR-----IVTAESDLNDPRVNVNPKKCGYNIDAQWVDFHHSIHAYLGER	323	
DB	332	FLAKLQNTWREAPHVMVTAEDSTEQVGTGAVEGGGLGFHYK-----	375	
QY	324	QGYTDFGNLDIVKSYKDFVYDGYKSNFRKTHGEPVGGELDCGNFVYIQ-----	375	
DB	376	-----NHQVQ-----NRKGERILIKLVDRSYKIAAALYLLSPYIPMFMGEEYEE	423	
QY	376	-----NHQVQ-----NRKGERILIKLVDRSYKIAAALYLLSPYIPMFMGEEYEE	423	
DB	413	FVLPPFSDHEVYVYKSLNKKMGPDYQWQKFA---QYRLLLGYMTVHPGKKLLFMGSEFAQ-	468	
QY	424	NPEYFSDFSKSLIQGVREGKKENGQDTPDQ-----ESTFNASKLSW---	468	
DB	469	-----FDEKNDTEQLDFWFLDSPMHQKASVFTQDLLRFYQKSKILYEHDRHQSFWDV	523	
QY	469	-KDEIFSFYKILMKRKLSEIADRRVNVVNGENWLLIKGREYFSLYVFSKSIEVYK	527	
DB	524	HNDEQISFIRYQKHGEALYIICNFTPVVYHQYD---VGVPFPTQYIEVLNSDSEY	579	
QY	528	SGT 530		
DB	580	GS 582		
RESULT	8			
GLGB_BACST	STANDARD;	PRT;	639	AA.
ID	GLGB_BACST			
AC	P30538:			
DT	01-APR-1993 (Rel. 25, Created)			
DT	01-APR-1993 (Rel. 25, Last sequence update)			
DE	01-FEB-1995 (Rel. 31, Last annotation update)			
DE	1,4-alpha-glucan branching enzyme (EC 2.4.1.18) (Glycogen branching enzyme).			
DE	enzyme).			
GN	GLGB.			
OS	Bacillus stearothermophilus			
OC	Bacteria; Firmicutes; Bacillus/Clostridium group;			
OC	Bacillus/Staphylococcus group; Geobacillus.			
OX	NCBI_TaxID=1422;			
RN	SEQUENCE FROM N.A.			
RP	STRAIN=1503-4R;			
RX	MEDLINE=92079888; PubMed=1745226;			
RT	Kiel J.A.K.W., Boels J.M., Beldman G., Venema G.;			
RT	"Molecular cloning and nucleotide sequence of the glycogen branching			
RT	enzyme gene (glgB) from Bacillus stearothermophilus and expression in			
RT	Escherichia coli and Bacillus subtilis."			

```
RESULT 9
ID GLGB_HAEIN STANDARD; PRT; 730 AA.
AC F45177;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 1.4-alpha-glucan branching enzyme (EC 2.4.1.18) (Glycogen branching
DE enzyme).
DE
DE
GN GLGB OR H11357.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RD / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus
RT influenzae Rd.";
RL Science 269:496-512(1995).
CC -!- CATALYTIC ACTIVITY: Formation of 1,6-glucosidic linkages of
CC glycogen.
CC
CC -!- PATHWAY: THIRD STEP IN GLYCOGEN BIOSYNTHESIS.
CC
CC -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC
CC -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U32815; AAC23004.1; -
CC TIGR; H11357; -
CC InterPro: IPR000461; Alpha-amylase.
CC DR InterPro: IPR004193; isoamylase_N.
CC DR Pfam: PF00128; alpha-amylase_1.
CC DR Pfam: PF02922; isoamylase_N_1.
CC KW Glycogen biosynthesis; transferase; Glycosyltransferase;
CC Complete proteome.
CC FT ACT_SITE 405 405 BY SIMILARITY.
CC FT ACT_SITE 458 458 BY SIMILARITY.
CC FT ACT_SITE 526 526 BY SIMILARITY.
CC SEQUENCE 730 AA; 83820 MW; 5B9575317F53769A CRC64;
CC
CC
CC Query Match 11.4%; Score 340; DB 1; Length 730;
CC Best Local Similarity 24.1%; Pred. No. 6.5e-15;
CC Matches 150; Conservative 101; Mismatches 214; Indels 158; Gaps 30;
CC
CC
CC QY 8 DG-NEVITLWAPYQKSVKLVLEKGLYMER-----DEKGYFTITLNNKVRDRYKV 60
CC ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
CC 133 DGVSGVNFRLWAPNARRVSI-VGDFNVWDRGRRHPRMFRHSKSGVWELFLPRASLGQLYKFE 191
CC
CC
CC QY 61 LDD-----ASLI-PDPASRYQPEGVHGPSQIIQSKENNETFLKKEDLII 105
CC ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
CC 192 LIDCHGNLRLKADFPFAFSQRLRPTAS--QVSALPNVVENTEARKKANQ-----GNOPISI 245
CC
CC
CC 106 YEIHVGTF--TPEGTF-----EGVIRKLDYLKDIGITAIEIMPIAOPGPKRDMGVDGYVL 158
CC ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
CC 246 YEVHLSGRRNLENNFWLDYDQIADELIPVVKEMGFTHIEFLPLSEFPDGSNGYQPLGL 305
CC
CC
CC QY 159 YAVONSYGSGPEGFKLVDEAHKGLGVLDVYNHVGPEGNVWVKGLPYFSQYKTP--- 215
CC ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
CC 306 YSPTSFRGSPFAFRLVKRAHEAGINVLDWVPGHFPDTHGLVAFDGTALYEHEDPREG 365
CC ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
CC
CC QY 216 ---WG-LTFNFDADSEVRKFFLENVYKWEYNVDGFRLLDAVHAIL----- 259
CC ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
CC 366 YHODWNTLIYNG---RNEVKNFLSSNALYLWLFVGVDGIRVDASMIYRDISRAGEW 422
CC
CC
CC QY 260 ---DTSFKHILEETADVH-----KYNRIVIAESDLNDPRVYNPKERCYNIDAQWV 308
CC ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
CC 423 IPNOYGGRENLEALEFLKHTNWKIHSMAGAISIAESTSFAGVTHPSENGGLGFNFKWN 482
CC ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
CC
CC QY 309 DDFHSHAYLTGERQGYTDFGNLDDIVKSY---KDFVYDGKYSNFRKKTGHEPVGEL 365
CC ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
CC 483 MGWMDTLAYN-----KLDPFYRQYHNKMTFGWVYQYSE----- 517
CC ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
CC
CC QY 366 DGCNFVYVYIIONHDOVGNRGKGERIILKLVDRSEYKIA--AALYLLS-----PYIPMFGEY 420
CC ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
CC 518 ---NFVLPL-SHDEVH-GKYSLLGKMPGDTWOKFANLRAYGYMGCYPOGKKLLFMGNEF 572
CC ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
CC
CC QY 421 GEENPEYFFSDFSLSKLIQVREGKKENGQDTPQDESTFNASKLSW-KIDBEI-PSFY 478
CC ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
CC 573 A-----QG-REWNVEES-----LDWFLDENIGGGWH 598
CC ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
CC
CC QY 479 KILIKRKELSIACDR-----VNVVNGENWLIIRKREYFSLYVFSKYSIEVYSGTLL 532
CC ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
CC 599 KGVLLKVLKDLNQYQKNRPLFELDNSPEGDFLWVDDAAN-SVLAFFERSNGE--RIIV 655
CC ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
CC
CC QY 533 LSSNNSFPQHIEE-----GKYE 549
CC ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
CC 656 VSNFTVPVRHNYRIGVNVAGKYE 678
CC ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
CC
CC RESULT 10
CC GLGB_STRAU STANDARD; PRT; 764 AA.
CC
CC AC F52980;
CC DT 01-OCT-1996 (Rel. 34, Created)
CC DT 01-OCT-1996 (Rel. 34, Last sequence update)
CC DT 01-OCT-1996 (Rel. 34, Last annotation update)
CC DE 1.4-alpha-glucan branching enzyme (EC 2.4.1.18) (Glycogen branching
CC enzyme).
CC DE
CC GN GLGB.
CC OS Streptomyces aureofaciens.
CC OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
CC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
CC OX NCBI_TaxID=1894;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN=CCM 3239 / ATCC 10762;
CC RX MEDLINE=94347823; PubMed=8068720;
CC RA Homerova D., Kormanec J.;
CC RT "Cloning of the putative glycogen branching enzyme gene, glgB, from
CC Streptomyces aureofaciens";
CC RL Biochim. Biophys. Acta 1200:334-336(1994).
CC CC -!- CATALYTIC ACTIVITY: Formation of 1,6-glucosidic linkages of
CC glycogen.
CC CC -!- PATHWAY: THIRD STEP IN GLYCOGEN BIOSYNTHESIS.
CC CC -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC CC -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC
```

```
CC EMBL; L11647; AAA67437.1; -
DR InterPro; IPR000461; Alpha_amylase.
DR InterPro; IPR001899; Gram_pos_anchor.
DR InterPro; IPR004193; isoamylase_N.
DR Pfam; PF00128; alpha-amylase; 1.
DR Pfam; PF02922; isoamylase_N; 1.
KW Glycogen biosynthesis; Transferase; Glycosyltransferase.
FT ACT_SITE 440 440 BY SIMILARITY.
FT ACT_SITE 493 493 BY SIMILARITY.
FT ACT_SITE 561 561 BY SIMILARITY.
SQ SEQUENCE 764 AA; 85325 MW; 6B45482E4A268ACF CRC64;

Query Match 11.2%; Score 335; DB 1; Length 764;
Best Local Similarity 24.2%; Pred. No. 1.5e-14;
Matches 133; Conservative 90; Mismatches 137; Indels 130; Gaps 26;

Qy 14 FTLWAPYQKSVKL---KVLEKGLYEMER-DEKGYFTITLNNVVRDRIKYL---DDAS 65
Dy 178 FTVWAPNALGVYRVGTGDFSYWDVAVYPMKSLGASGWELFLPGVAGALYKYEITRPDGR 237
Qy 66 EI-DPPASRYQPEGVHGVSQIQESKEFNNETFLKKEDLI-----IYEHVGTFTPE 116
Dy 238 TLRADPMARYAEVPPANASIVTASRYEWDREWARRCALAPHAQPMVSVELHLASWRPG 297
Qy 117 GTFEGVIRKLD-YLKDLGITAIEIMPIAOPFGKRDWGDYVLYAVQNSYGGPEGFRKL 175
Dy 298 LSYRQLAEQLPAYVKELGFTHVELMPVAEHPFGGSGVQVTFGYAPTSRMGTDPDFRLV 357
Qy 176 DEAKKGLGVLDVYVNVHVGPEGNYVWKLGPYFSQKYKTP-----WGLTFNDDAESD 228
Dy 358 DALHRAGIGVIVDVPVAPHPFRDDWALAEFDGRPLYEHQDPRRAHPDWG-TLEFDYGRK 415
Qy 229 EVRKFILENVYWKYKENVNDCGRFLDAVHAI-----DTSPK-----HTLE 268
Dy 416 EVRNLVANAYWCQEFHVDGLRADAVASMLYDYSRDEGDSWSPNAGGREDLDAVALLQ 475
Qy 269 EI-ADVHKYNRIV-IAESLDNPRVNPKEKCGYNIDQAVDDFHHSIAHYLGE---R 323
Dy 476 EMNATVYRRFPGVVTIAEESTAWDGVTRPTDSGGGLGKWNMGWMDTLRYVSKPEVHR 535
Qy 324 QGYTID-FGNLDDIVKSYKDVYVDGKYSNFRKTHGEPVGLDGCNFFVYVIONHDOVG 381
Dy 536 KYHHDMTFG---WVAFSE-----NFVLPV-SHDEV 564
Qy 382 NRGKGERIKL-----VDRESYKIAAALYLLSPYIPMFMEGEY--GEE-----NPYFF 429
Dy 565 H-GRSLVSKMPGDWQWQORATHRAYLGFMAHPGKQLLFMGQEFQAQGSSESETYGPDWV 623
Qy 430 SDFS-----DSKLIQGVREGKKENGQDTPQDESTFNASKLSWK----- 469
Dy 624 LDSSYPAGDHF---LGVRSVLRLN-----RTVTASPALWERDSVPEGFARWEADA 671
Qy 470 IDEEIFSFK 479
Dy 672 ADDNVFAFLR 681

RESULT 11
GLGB_SYNY3
ID GLGB_SYNY3 STANDARD; PRT; 770 AA.
AC P52981;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 1,4-alpha-glucan branching enzyme (EC 2.4.1.18) (Glycogen branching
DE enzyme).
GN GLGB OR SLL0158.
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1148;
RN [1]
```

```
RP SEQUENCE FROM N.A.
RX MEDLINE=96127529; PubMed=8590279;
RA Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,
RA Sugitara M., Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
RT region from map positions 64% to 92% of the genome.";
RL DNA Res. 2:153-166(1995).
CC -!- CATALYTIC ACTIVITY: Formation of 1,6-glucosidic linkages of
CC glycogen.
CC -!- PATHWAY: THIRD STEP IN GLYCOGEN BIOSYNTHESIS.
CC -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D63999; BAA10073.1; -
DR InterPro; IPR000461; Alpha_amylase.
DR InterPro; IPR004193; isoamylase_N.
DR Pfam; PF00128; alpha-amylase; 1.
DR Pfam; PF02922; isoamylase_N; 2.
KW Glycogen biosynthesis; Transferase; Glycosyltransferase;
KW Complete proteome.
FT ACT_SITE 433 433 BY SIMILARITY.
FT ACT_SITE 486 486 BY SIMILARITY.
FT ACT_SITE 554 554 BY SIMILARITY.
SQ SEQUENCE 770 AA; 89527 MW; A435AFCA7703FABA CRC64;

Query Match 11.2%; Score 333.5; DB 1; Length 770;
Best Local Similarity 23.1%; Pred. No. 1.9e-14;
Matches 119; Conservative 91; Mismatches 176; Indels 129; Gaps 20;

Qy 7 IDG-NEVIFTLWAPYQKSVKL---KVLEKGLYEMERDEKGYFTITLNNVVRDRIKYL 61
Dy 136 VDGKGYFFAWAPNARNVSIILGDFNNDGLHOMRRKNNWELFPELGVGTSYKYEI 195
Qy 62 DD-----ASEIPDPSRYQPEGVHGVSQIIQ-ESKEFNNETFLK-----KEDLIYEI 108
Dy 196 KNEGHIYEKTDPGYQYQEVPRKTAIVADLDGQWHDWLEARRTSDPLSKPVSVEL 255
Qy 109 HVG-----TFTPECT-----FEGVIRKLDYKLDGITAIEIM 140
Dy 256 HLGSLWHTAYDEPVKTLHGEGVPVEVSEWNTGARFLTYEYELVDKLIIPYKELGYTHIEL 315
Qy 141 PIAQFPCKRDMGYGVLYAVONSYGGPEGPRKLVDEAHKGLGVILDVVYNHVGPEGN- 199
Dy 316 PIAEHPDGSWGKYQVYGYAPTSPFGSPEDPMYFVDOCHLNGIIVIDWPCHPKDGHG 375
Qy 200 -----YMWKLGYPYFSQKYKTPMG-LTFNPDAADESDEVRKFFILENVEYWKYKENVDFGR 251
Dy 376 LAFFDGTGLYEHGDPKRGHEK-ENGTLIFNYG---RNEVRNLFVANALFWFDKYHIDGMR 431
Qy 252 LDVAHAI-----DTSPKHILSEIADVVHKY--NRIVIAESLDN 289
Dy 432 VDAVASMLYLDYCREGEWANEYGGRENLEAADFLLQVNSVWYSYFPGLISIAESTSW 491
Qy 290 PRVNPKEKCGYNIDQAVDDFHHSIAHYLGTGERQYVTFDGNLDDIVKSYKDVFFVDGK 349
Dy 492 PWSWPTVYVGLGNLKW-----NWGMHMDL-----DFFSMDPW 526
Qy 350 YSNFRKT-----HGEVPGELDGCNFFVYVIONHDOVGNRGKGERIKLVDRESYKIA 401
Dy 527 FRQPHQNSITFSMWYNHSE-----NYMLAL-SHDEVVH-GKSNMLGKMPGDEWKYA 576
Qy 402 AALYLLS-----PYIPMFMEGEYGEENPFYFFSD 431
```


Db 577 NVRALFTYMTHTPKKTKTMSMEFGQWSENNVWGD 611
 RESULT 12
 GLGB_SYN7 STANDARD; PRT; 773 AA.
 ID GLGB_SYN7 STANDARD; PRT; 773 AA.
 AC P16934;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 01-FEB-1995 (Rel. 31, Last annotation update)
 DE 1,4-alpha-glucan branching enzyme (EC 2.4.1.18) (Glycogen branching
 enzyme).
 DE enzyme).
 GN GLGB.
 OS Synecococcus sp. (strain PCC 7942) (Anacystis nidulans R2).
 OC Bacteria; Cyanobacteria; Chroococcales; Synecococcus.
 OC NCBI_TaxID=1140;
 RN [1]
 RP MEDLINE=90323609; PubMed=2142668;
 RX Kiel J.A.K.W., Boels J.M., Beldman G., Venema G.;
 RT "Nucleotide sequence of the Synecococcus sp. PCC7942 branching
 enzyme gene (glbB): expression in Bacillus subtilis.";
 RL Gene 89:77-84(1990).
 CC -1- CATALYTIC ACTIVITY: Formation of 1,6-glucosidic linkages of
 glycogen.
 CC -1- PATHWAY: THIRD STEP IN GLYCOGEN BIOSYNTHESIS.
 CC -1- SUBUNIT: MONOMER.
 CC -1- MISCELLANEOUS: THE TEMPERATURE FOR OPTIMAL ACTIVITY IS
 APPROXIMATELY 35 DEGREES CELSIUS.
 CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
 KNOWN AS THE ALPHA-AMYLASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M31544; AAB39038.1; -
 DR PIR: JQ0550.
 DR InterPro: IPR000461; Alpha_amylase.
 DR InterPro: IPR004193; isoamylase_N.
 DR Pfam: PF00128; alpha-amylase; 1.
 DR Pfam: PF02922; isoamylase_N; 1.
 KW Glycogen biosynthesis; Transferase; Glycosyltransferase.
 FT INIT_MET 0 0
 FT ACT_SITE 439 439 BY SIMILARITY.
 FT ACT_SITE 492 492 BY SIMILARITY.
 FT ACT_SITE 560 560 BY SIMILARITY.
 FT SEQUENCE 773 AA; 89063 MW; 52BAA17CA337BF57 CRC64;
 Query Match 11.0%; Score 326.5; DB 1; Length 773;
 Best Local Similarity 24.2%; Pred. No. 5.3e-14;
 Matches 121; Conservative 78; Mismatches 175; Indels 125; Gaps 20;
 QY 12 VIFLWAPYQSKVL-----KVLKGLYMERDEKGYFTTLNKKVDRYKVLDDAS-- 65
 Db 147 VNFVAPARNVNSILGDFNSWDGRKHQMARNSGIWELFPELTVGAAKYKKNVDGH 206
 QY 66 --EIPDPASRYOPEGVHGPSQIOESKEP--NNEFTLK-----KEDLIYETHGCTF 113
 Db 207 IYEKSDPYG-FQOEVRPKTASIVADLDRTYWGDAWLERRRHQEPDPISVYEVHLSW 265
 QY 114 -----TPECTF-----EGVIRKLDYKLGKLGTAIEIMPTAQ 144
 Db 266 MHASSDAIATDAQKPLPPVPVADLKPGARFLTYRELADRLIPYVLDLGYSHIELLPAT 325
 QY 145 FPKCRDWDGYVLYAVQNSYGGPGKFLVDHAKHKGVLVDVYVNHVGPGEYMYVKL 204

```
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z73902; CAA98090.1; -.
DR EMBL: AE007010; AAK45632.1; -.
DR TIGR: MT1368; -.
DR InterPro: RVI326c; -.
DR InterPro: IPR000461; Alpha_amylase.
DR InterPro: IPR004193; isoamylase_N.
DR Pfam: PF00128; alpha-amylose; 1.
DR Pfam: PF02922; isoamylase_N; 1.
KW Glycogen biosynthesis; Transferase; Glycosyltransferase;
KW Complete proteome.
FT ACT_SITE 411 411 BY SIMILARITY.
FT ACT_SITE 464 464 BY SIMILARITY.
FT ACT_SITE 532 532 BY SIMILARITY.
FT CONFLICT 214 214 P -> A (IN REF. 2).
FT CONFLICT 223 223 P -> Q (IN REF. 2).
SQ SEQUENCE 731 AA; 81729 MW; BE2BFEF765352617 CRC64;

Query Match 10.98; Score 326; DB 1; Length 731;
Best Local Similarity 22.88; Pred. No. 5.4e-14;
Matches 117; Conservative 76; Mismatches 186; Indels 134; Gaps 18;

QY 12 VIFTLWAPYQKSVKL-----KVL-EKGLYEMERDE---KGYFTITLNNVK- 52
DB 147 VSFVAPNAKGVSLIGFNCWNGHEAPRMVLGSPGVWELFWPDPDGLYKFRVHGADG 206
QY 53 -VRORYKVLDDASEIFDPASRYOPEGVHGSPSQIIQESKEFNFTFLKEDLIIEIHVG 111
DB 207 VVTDRADPFAPGTEVPQTASRVTSSTDTWGDWDMAGRLNPV---NEAMSTYEVHLG 263
QY 112 TFTPCTEGVIRKL-DYLDGLGITATEIMPIAQFPCKRDWGYDGVLYAVQNSYGGPEG 170
DB 264 SWRPGLSYQLARELTDYIVDQGFTHVLLPVAEHPFAGSGYQVTSYTAPTSFRFGTPDD 323
QY 171 FRKLVDRAHKKGLGVLDV-----YNHVGPENGMVYKLGPPYSQK 211
DB 324 FRALVDALHQAGIGVVDWVPAHPKDAWALGRDGTPLYEHSDPKRGELDWGTY---- 379
QY 212 YKTPWGLTFNPDDAESDEVKRFLENYVEIKYKENVGDFRLDAVHAI----- 259
DB 380 -----VFDFG---RPEVRNLFVANALYWLQEPHIDGLRVDASVSMYLDYSRPEGWT 429
QY 260 -----DTSPKHILEEADVVHKY--NRIVIAESDLNDPRVVPNPKKCGYNIDAQWVD 309
DB 430 PNVHGGRENLEAVQFQEMNATAHKVAPGIVTIAESTPWSGVTRPTNIGGLGFSMKWM 489
QY 310 DFHHSIHAYLTGERQYITDFGNLDDIVKSYKD-----VFVYDGYKSNFRKRTHGEPV 362
DB 490 GWMHDT-----LDVSRDPYRSYHHHEWTFMSLYAFSENY----- 525
QY 363 GELDGCNFVYIQNDQVNGRKGRIKILVDRESYKIAAALYLLS-----PYIPMIFMG 417
DB 526 -----VLPLSHDEVVH--GKTLGRMPGNHNNHKAAGLSLLAYQAWHPGKQLLFMG 575
QY 418 EEVGEENPFYFFSDSKLIQVREGKCKENG 450
DB 576 QEFQGR-----AEWSEQR---GLDWFQDENG 599

RESULT 14
ID GLGB_BUTFI
AC P30539;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)
DE 1.4-alpha-glucan branching enzyme (EC 2.4.1.18) (Glycogen branching
DE enzyme).
```

```
GN GLGB.
OS Butyrivibrio fibrisolvens.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Butyrivibrio.
OX NCBI_TaxID=831;
RN [1]
RC SEQUENCE FROM N.A., AND SEQUENCE OF 1-12.
RP STRAIN=H17C;
RX MEDLINE=92041554; PubMed=1938880;
RA Rumbak E., Rawlings D.E., Lindsey G.G., Woods D.R.;
RT "Characterization of the Butyrivibrio fibrisolvens glgB gene, which
RT encodes a glycogen-branching enzyme with starch-clearing activity.";
RL J. Bacteriol. 173:6732-6741(1991).
CC -I- FUNCTION: TRANSFERS CHAINS OF 5 TO 10 (OPTIMUM, 7) GLUCOSE UNITS,
CC USING AMYLOSE AND AMYLOPECTIN AS SUBSTRATES, TO PRODUCE A HIGHLY
CC BRANCHED POLYMER.
CC -I- CATALYTIC ACTIVITY: Formation of 1,6-glucosidic linkages of
CC glycogen.
CC -I- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M64980; AAA23007.1; -.
DR PIR: B41328; B41328.
DR InterPro: IPR000461; Alpha_amylase.
DR InterPro: IPR004193; isoamylase_N.
DR Pfam: PF00128; alpha-amylose; 1.
DR Pfam: PF02922; isoamylase_N; 1.
KW Glycogen biosynthesis; Transferase; Glycosyltransferase.
FT ACT_SITE 313 313 BY SIMILARITY.
FT ACT_SITE 366 366 BY SIMILARITY.
FT ACT_SITE 434 434 BY SIMILARITY.
SQ SEQUENCE 639 AA; 73875 MW; DC93B72A5F28C2AF CRC64;

Query Match 10.88; Score 321; DB 1; Length 639;
Best Local Similarity 25.18; Pred. No. 9.5e-14;
Matches 124; Conservative 81; Mismatches 153; Indels 136; Gaps 27;

QY 14 FTLWAPYQKSVKL---KVLKGLYEMERDEKG-YFTITLNNYKVRDXYKVL--DDASE 66
DB 42 FAVWAPNAADVHVVDGDFNGWDENAHQMKRSKTGNITWLFIPGVAIGALYKFLTAQDGRK 101
QY 67 I--PDPASRY---QPEGVHGSPSQII---QESKEFNNETFLKKED-----LIIEIHVGT 112
DB 102 LYKADPANYAEALPGNASRTDLSGFKWSDSKWYES---LKGDMNRQPIAYIECHIGS 158
QY 113 FT--PEGTFFGVI-----RKLDYLKDLGITATEIMPIAQFPCKRDWGYDGVLYAVQN 163
DB 159 WMKHPDGTEDGFYTYRQFADRIVEYKEMKYTHIELIGIAEHHPDGSWGQVGTGYAAPT 218
QY 164 SYGSGPEGRKLVDEAHKKGLGVLDVYVNVGP-----EGNYMVKLGPYSQKYKTP 215
DB 219 RYGEPTDFMYLINQLHKKHGIGVILDVWVPAHPFCDFEGLACFDGTCI-----YEDP 268
QY 216 -----WGL-TFNFDDAESDEVKRFLENYVEIKYKENVGDFRLDAVHAI--IDTSP 263
DB 269 DPRKGEHPDWGTKIFNL---AKPEVKNFLIANALYWRKPHIDGLRVDASVSMYLDYDK 325
QY 264 K-----HILEEADVVHKYNRIV-----TAESDLNDPRVVPNPKKCGYN 302
DB 326 KGGQWVPNKYGDKNLDAIEFFKH-FNSVVRGTGYPTNLTIAEESTAMPKVTAPPEEDGLG 384
QY 303 IDAQWDDFHSHIAYLTG--RQ--YYTDFGNLDDIVKSYKDVVYDGYKSNFRKRT 357
DB 385 FAFKWNMGWHDCEYMKLDYPRQGAHYWMTF-----AMSYN----- 432
```

```
QY 358 HGEFVGLDGCNEFVYQIONHDOVGNRGKGRRIIKLVDRSYKIAAALYLLSPYIPM----- 413
Db 423 -----DSENVILPL-SHDEVVHL-KCSMVEKM---PGYKVDKYANLRVGYTYMRGHS 469

QY 414 -----IFWGEYBEE 423
Db 470 GKLLFMGQDFGQE 483

RESULT 15
GLGX_HAEN STANDARD; PRT; 659 AA.
AC P45178;
DT 01-NOV-1995 (rel. 32, Created)
DT 16-OCT-2001 (rel. 32, Last sequence update)
DE Glycogen operon protein glgx (EC 3.2.1.-).
GN GLGX OR H11358.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RD / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uterback G.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghegan N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus
RT influenzae Rd.";
RL Science 269:496-512(1995).
CC -!- FUNCTION: THIS PROTEIN MAY BE PART OF A GLYCOGEN BIOSYNTHETIC/
CC CATABOLIC OPERON BUT IS NOT REQUIRED FOR GLYCOGEN SYNTHESIS
CC (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U32815; AAC23005.1; -
DR HSSP; P10342; 1BF2.
DR TIGR; H11358; -
DR InterPro; IPR000461; Alpha-amylase.
DR InterPro; IPR004193; isoamylase_N.
DR Pfam; PF00128; alpha-amylase_1.
DR Pfam; PF02922; isoamylase_N; 1.
KW Glycogen biosynthesis; Hydrolase; Glycosidase; Complete proteome.
FT ACT_SITE 333 333 BY SIMILARITY.
FT ACT_SITE 440 440 BY SIMILARITY.
SQ SEQUENCE 659 AA; 75290 MW; E2B71F610E8B1CAC CRC64;

Query Match 10.7%; Score 318; DB 1; Length 659;
Best Local Similarity 23.1%; Pred. No. 1.6e-13;
Matches 137; Conservative 98; Mismatches 173; Indels 186; Gaps 34;

QY 3 FAYKIDG-----NEVIFTWAPYQKSVKLVLEKGLYEMERDE-KGYFTITLNNVKVRDR 56
Db 74 YAFRIHGEFANQKILD----PYAKAV-----NGKPDLSSESRSWFLLSDN-----RDN 120
```

```
QY 57 YKVLDDASEIPDAPRYOQEPGVHGPSQIIQESKEFNNETFLKK--EDLIYIHHVGTFT 114
Db 121 -----AHLAPRAV-----VISEEDFENDSPNTPAETIYVELHWKGF 160

QY 115 P-----EGTFEGVIR--KLDYKDLGITAIEIMPI-----AOPPGKRD-WGYD 154
Db 161 QLNKEIPAALRGTYTGLAHPVNLAYLKLGLGTAVELLPVNFHNEPHLQARGLQNYWGN 220

QY 155 GVVLYAVONSYGPEG-----FRKLVDEAHKKGVLVDVYVNH-VGPEGNMYVKLGYPFS 209
Db 221 PLAMFAVEPKYATNPNPLAEFTKTMVKAFAHKGAGIEVLVDVVFVNSAESEQTY-----PTFS 275

QY 210 QK-----YKTPWGLTFNFD-----DAESDEVKRFLENYVEYWKYKYNVDGFRLDV 255
Db 276 QRGIDQOQTYWRNDQGRYINWTCCGNMLNLSDDVGRKWVVDCLRYVWEQCHIDGFRDLA 335

QY 256 HAIDTSPKHILEEADVVHKYNRIVIAESDL-NDRPVVNPK-----EKGYNID-- 304
Db 336 TVLGRDTP-----DFNSSAQLFTDIKNPSLQNIKLIABPMDIGHYGYQVGNF 383

QY 305 ----AQVDDDFHHSI---HAYLTGERQGYTDFGNLDDIVK----- 338
Db 384 PSYFAEWNDRFRDDLCRFWLWKSGETGAFAERFAGSSDLFKKNDRLPHTLTNITAHDFG 443

QY 339 SYKDVVFYDGYKSYNFRKRTHGEPVGLDGCNFVYIYQNHDOVGNRGKGERIIR-----LV 393
Db 444 TLKDLVSYNQKH-----ETNGEE--NRDGRN-ENYSYNH---GVEGSTESLSEPKSAVE 493

QY 394 DRESYKIAAAL--YLLSPYIPMIFMGEEYGEENPFYFFSDFSKLIQGVREGKKENGQ 451
Db 494 NNRTFAQSGLLSLLLANGTPLLADGDFG-----NFTQGNNNAYCQ 535

QY 452 DTDPDQDESTFNASKLSW-KIDEEIFSYKTLIKMKKELSIACDRRVNVVNGENW 504
Db 536 D-----NEITW-----LKWANFNEELFELTKQTIALRKO-----IGSLUNKDQW 573
```

Search completed: July 15, 2002, 12:24:32
Job time: 384 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 12:17:38 ; Search time 55.94 Seconds
(without alignments)
1728.712 Million cell updates/sec

Title: US-09-298-924-6
Perfect score: 2978
Sequence: 1 MTFAYKIDGNEVFTLWAPY.....PQHIEGKYEFDFGKGFALYKL 559

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : SPTREMBL_19.*
- 1: sp_archaea.*
 - 2: sp_bacteria.*
 - 3: sp_fungi.*
 - 4: sp_human.*
 - 5: sp_invertebrate.*
 - 6: sp_mammal.*
 - 7: sp_mhc.*
 - 8: sp_organelle.*
 - 9: sp_phase.*
 - 10: sp_plant.*
 - 11: sp_rodent.*
 - 12: sp_virus.*
 - 13: sp_vertebrate.*
 - 14: sp_unclassified.*
 - 15: sp_rvirus.*
 - 16: sp_bacteriap.*
 - 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2978	100.0	559	1 Q55088	Q55088 sulfolobus
2	2837	95.3	559	1 Q9UWN9	Q9UWN9 sulfolobus
3	2467	82.8	561	17 P95867	P95867 sulfolobus
4	1748.5	58.7	556	1 Q53641	Q53641 sulfolobus
5	1269	42.6	359	17 Q973H4	Q973H4 sulfolobus
6	1033.5	34.7	581	2 Q9AD15	Q9AD15 streptomyces
7	905.5	30.4	600	16 Q9RX51	Q9RX51 deinococcus
8	881.5	29.6	583	16 Q911V1	Q911V1 pseudomonas
9	876.5	29.4	601	16 Q92U63	Q92U63 rhizobium m
10	706	23.7	217	17 Q973H5	Q973H5 sulfolobus
11	524.5	17.6	552	2 Q44528	Q44528 anabaena va
12	399	13.4	652	2 Q59242	Q59242 bacillus st
13	397	13.3	440	2 Q93Q35	Q93Q35 myxococcus
14	395.5	13.3	1142	2 Q93029	Q93029 bacillus sp
15	381.5	12.8	825	2 Q59319	Q59319 caldocellum
16	365	12.3	1280	16 Q97SQ7	Q97SQ7 streptococc

17	364	12.2	1287	2	Q9F930	Q9F930 streptococc
18	361	12.1	630	16	O66936	O66936 aquifex aeo
19	349.5	11.7	783	10	O04196	O04196 arabidopsis
20	348.5	11.7	666	16	O84046	O84046 chlamydia t
21	346.5	11.6	713	1	O05152	O05152 sulfolobus
22	345.5	11.6	422	2	O59243	O59243 bacillus st
23	345.5	11.6	664	16	Q928F5	Q928F5 chlamydia p
24	345.5	11.6	666	16	Q9PKZ6	Q9PKZ6 chlamydia m
25	345.5	11.6	707	16	P73608	P73608 synechocyst
26	344.5	11.6	741	2	Q9KV06	Q9KV06 streptomyces
27	344.5	11.6	741	2	Q59832	Q59832 streptomyces
28	343	11.5	718	17	P95868	P95868 sulfolobus
29	341	11.5	737	16	Q985P4	Q985P4 rhizobium 1
30	339.5	11.4	789	10	O22637	O22637 zea mays (m
31	336	11.3	702	10	Q9M0S5	Q9M0S5 arabidopsis
32	336	11.3	716	17	Q973H3	Q973H3 sulfolobus
33	335	11.2	1072	16	Q9K6N1	Q9K6N1 bacillus ha
34	334.5	11.2	717	16	Q9K7U5	Q9K7U5 bacillus ha
35	334.5	11.2	736	16	Q92M14	Q92M14 rhizobium m
36	333.5	11.2	720	16	Q9RXP5	Q9RXP5 deinococcus
37	331.5	11.1	1938	2	P70983	P70983 bacillus sp
38	330.5	11.1	562	2	Q9L872	Q9L872 bacillus sp
39	330.5	11.1	562	16	Q9K7J1	Q9K7J1 vibrio chol
40	329.5	11.1	668	2	P71095	P71095 bacteroides
41	329.5	11.1	818	10	Q41742	Q41742 zea mays (m
42	329.5	11.1	1165	16	Q99XX8	Q99XX8 streptococc
43	329	11.0	741	3	Q9P5P3	Q9P5P3 neurospora
44	327	11.0	733	10	O80403	O80403 oryza sativ
45	323.5	10.9	718	16	O34587	O34587 bacillus su

ALIGNMENTS

RESULT	1
Q55088	PRELIMINARY; PRT; 559 AA.
AC	Q55088;
DT	01-NOV-1996 (TRENBLrel. 01, Created)
DT	01-MAY-1997 (TRENBLrel. 03, Last sequence update)
DT	01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE	ALPHA-AMYLASE.
OS	Sulfolobus solfataricus.
OC	Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
OX	NCBI_TaxID=2287;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=KMI;
RA	Kato M., Kettoku M., Miura Y., Komeda T., Konishi Y., Shindo K.,
RA	Kobayashi K., Iwanatsu A.;
RT	"The gene analysis of the new amylases from the hyper thermophilic
RT	archae Sulfolobus";
RL	Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.
DR	EMBL; D64130; BAAL1010.1; -.
DR	InterPro; IPR000461; Alpha-amylase.
DR	InterPro; IPR004193; isoamylase_N.
DR	Pfam; PF00128; alpha-amylase; 1.
DR	Pfam; PF02922; isoamylase_N; 1.
SQ	SEQUENCE 559 AA; 64790 MW; 79FBE23A7CD38B4E CRC64;

Query Match 100.0%; Score 2978; DB 1; Length 559;
Best Local Similarity 100.0%; Pred. No. 1.3e-180;
Matches 559; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MTFAYKIDGNEVFTLWAPYQKSVKLVLEKGLYEMERDEKGYFTITLNNKVRDRYKYV	60
Db	1	MTFAYKIDGNEVFTLWAPYQKSVKLVLEKGLYEMERDEKGYFTITLNNKVRDRYKYV	60
QY	61	LDASIPDPASRYQEGVHGPSQIIQESKEFNNEFLKEDLLIYEHVGTTPGCTFE	120
Db	61	LDASIPDPASRYQEGVHGPSQIIQESKEFNNEFLKEDLLIYEHVGTTPGCTFE	120

```
OY 121 GVIRKLDYKDLGITAIEIMPTAQPPGKRDWGYDGYLYAVONSYGGPGEFRKLVDEAHK 180
Db 121 GVIRKLDYKDLGITAIEIMPTAQPPGKRDWGYDGYLYAVONSYGGPGEFRKLVDEAHK 180
OY 181 KGLGVILDVYVNHVGPENYMWKLGYPYSQYKTPWGLTFNFDAAESDEVKRFILLENVEY 240
Db 181 KGLGVILDVYVNHVGPENYMWKLGYPYSQYKTPWGLTFNFDAAESDEVKRFILLENVEY 240
OY 241 WKEYNVDGFRDLAVHAIIDTSPKHILEEADVHVHKNRIVIAESDLNDPRVVPKKECG 300
Db 241 WKEYNVDGFRDLAVHAIIDTSPKHILEEADVHVHKNRIVIAESDLNDPRVVPKKECG 300
OY 301 YNIDAQWVDDFHHSIHAYLTGERQGYDYGNDLDDIVKSYKDVYDGYKSNFRKKTGCE 360
Db 301 YNIDAQWVDDFHHSIHAYLTGERQGYDYGNDLDDIVKSYKDVYDGYKSNFRKKTGCE 360
OY 361 PVGELDGCNFVYVYQNHQDVGNRGKGERIKLVDRSEYKIAAALYLLSPYIPMIFPMGEY 420
Db 361 PVGELDGCNFVYVYQNHQDVGNRGKGERIKLVDRSEYKIAAALYLLSPYIPMIFPMGEY 420
OY 421 GEENPFYFFSDSKLIQGVREGKKGQDTPDQDESTFNASKLSWKIDEEIFSFKYI 480
Db 421 GEENPFYFFSDSKLIQGVREGKKGQDTPDQDESTFNASKLSWKIDEEIFSFKYI 480
OY 481 LIKMKKELSIACDRRVNVVNGENWLLIKGREYFSLYVFSKSIEVKYSGTLLLSNNSPF 540
Db 481 LIKMKKELSIACDRRVNVVNGENWLLIKGREYFSLYVFSKSIEVKYSGTLLLSNNSPF 540
OY 541 QHIEGKYEFKGFALYKL 559
Db 541 QHIEGKYEFKGFALYKL 559

RESULT 2
OY 121 GVIRKLDYKDLGITAIEIMPTAQPPGKRDWGYDGYLYAVONSYGGPGEFRKLVDEAHK 180
Db 121 GVIRKLDYKDLGITAIEIMPTAQPPGKRDWGYDGYLYAVONSYGGPGEFRKLVDEAHK 180
OY 181 KGLGVILDVYVNHVGPENYMWKLGYPYSQYKTPWGLTFNFDAAESDEVKRFILLENVEY 240
Db 181 KGLGVILDVYVNHVGPENYMWKLGYPYSQYKTPWGLTFNFDAAESDEVKRFILLENVEY 240
OY 241 WKEYNVDGFRDLAVHAIIDTSPKHILEEADVHVHKNRIVIAESDLNDPRVVPKKECG 300
Db 241 WKEYNVDGFRDLAVHAIIDTSPKHILEEADVHVHKNRIVIAESDLNDPRVVPKKECG 300
OY 301 YNIDAQWVDDFHHSIHAYLTGERQGYDYGNDLDDIVKSYKDVYDGYKSNFRKKTGCE 360
Db 301 YNIDAQWVDDFHHSIHAYLTGERQGYDYGNDLDDIVKSYKDVYDGYKSNFRKKTGCE 360
OY 361 PVGELDGCNFVYVYQNHQDVGNRGKGERIKLVDRSEYKIAAALYLLSPYIPMIFPMGEY 420
Db 361 PVGELDGCNFVYVYQNHQDVGNRGKGERIKLVDRSEYKIAAALYLLSPYIPMIFPMGEY 420
OY 421 GEENPFYFFSDSKLIQGVREGKKGQDTPDQDESTFNASKLSWKIDEEIFSFKYI 480
Db 421 GEENPFYFFSDSKLIQGVREGKKGQDTPDQDESTFNASKLSWKIDEEIFSFKYI 480
OY 481 LIKMKKELSIACDRRVNVVNGENWLLIKGREYFSLYVFSKSIEVKYSGTLLLSNNSPF 540
Db 481 LIKMKKELSIACDRRVNVVNGENWLLIKGREYFSLYVFSKSIEVKYSGTLLLSNNSPF 540
OY 541 QHIEGKYEFKGFALYKL 559
Db 541 QHIEGKYEFKGFALYKL 559

RESULT 2
ID Q9UNW9 PRELIMINARY; PRT; 559 AA.
AC Q9UNW9;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE MALTOOLIGOSYL TREHALOSE TREHALOHYDROLASE.
GN TREZ.
OS Sulfolobus shibatae.
OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
OX NCBI_TaxID=2286;
RN [1]
RP SEQUENCE FROM N.A.
RC Weiting Y., Wei C., Hui W., Li L., Cheng J.;
RX "Maltoligosyl trehalose trehalohydrolase from Sulfolobus shibatae.";
RT Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RL EMBL; AF201335; AAF17553.1;
DR EMBL; AF201335; AAF17553.1;
DR InterPro; IPR000461; Alpha_amylase.
DR InterPro; IPR004193; Isoamylase_N.
DR Pfam; PF00128; alpha-amylase; 1.
DR Pfam; PF02922; isoamylase_N; 1.
KW Hydrolase.
SQ SEQUENCE 559 AA; 64619 MW; 8C0F0F38A5AE2846 CRC64;

Query Match 95.3%; Score 2837; DB 1; Length 559;
Best Local Similarity 95.2%; Pred. No. 1.1e-171;
Matches 532; Conservative 11; Mismatches 16; Indels 0; Gaps 0;

OY 1 MTFAYKIDGNEVIFTLWAPYQKSVKLVLEKGLYEMERDEKGYFTITLNNVKVDRYKYV 60
Db 1 MTFYKIDGNEVIFTLWAPYQKSVKLVLEKGLYEMERDDKGYFTITLNNVKVDRYKYV 60

OY 61 LDDASEIPDPASRYQPGVGPSPQIIQESKEFNNETLKKEDLIYIHHVGTFTPEGTTFE 120
Db 61 LDDASEIPDPASRYQPGVGPSPQIIQESKEFNNETLKKEDLIYIHHVGTFTPEGTTFE 120

OY 121 GVIRKLDYKDLGITAIEIMPTAQPPGKRDWGYDGYLYAVONSYGGPGEFRKLVDEAHK 180
Db 121 GVIRKLDYKDLGITAIEIMPTAQPPGKRDWGYDGYLYAVONSYGGPGEFRKLVDEAHK 180
```

```
Db 121 GVIRKLDYKDLGITAIEIMPTAQPPGKRDWGYDGYLYAVONSYGGPGEFRKLVDEAHK 180
OY 181 KGLGVILDVYVNHVGPENYMWKLGYPYSQYKTPWGLTFNFDAAESDEVKRFILLENVEY 240
Db 181 KGLGVILDVYVNHVGPENYMWKLGYPYSQYKTPWGLTFNFDAAESDEVKRFILLENVEY 240
OY 241 WKEYNVDGFRDLAVHAIIDTSPKHILEEADVHVHKNRIVIAESDLNDPRVVPKKECG 300
Db 241 WKEYNVDGFRDLAVHAIIDTSPKHILEEADVHVHKNRIVIAESDLNDPRVVPKKECG 300
OY 301 YNIDAQWVDDFHHSIHAYLTGERQGYDYGNDLDDIVKSYKDVYDGYKSNFRKKTGCE 360
Db 301 YNIDAQWVDDFHHSIHAYLTGERQGYDYGNDLDDIVKSYKDVYDGYKSNFRKKTGCE 360
OY 361 PVGELDGCNFVYVYQNHQDVGNRGKGERIKLVDRSEYKIAAALYLLSPYIPMIFPMGEY 420
Db 361 PVGELDGCNFVYVYQNHQDVGNRGKGERIKLVDRSEYKIAAALYLLSPYIPMIFPMGEY 420
OY 421 GEENPFYFFSDSKLIQGVREGKKGQDTPDQDESTFNASKLSWKIDEEIFSFKYI 480
Db 421 GEENPFYFFSDSKLIQGVREGKKGQDTPDQDESTFNASKLSWKIDEEIFSFKYI 480
OY 481 LIKMKKELSIACDRRVNVVNGENWLLIKGREYFSLYVFSKSIEVKYSGTLLLSNNSPF 540
Db 481 LIKMKKELSIACDRRVNVVNGENWLLIKGREYFSLYVFSKSIEVKYSGTLLLSNNSPF 540
OY 541 QHIEGKYEFKGFALYKL 559
Db 541 QHIEGKYEFKGFALYKL 559

RESULT 3
ID P95867 PRELIMINARY; PRT; 561 AA.
AC P95867;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ALPHA-AMYLASE PRECURSOR (MALTO-OLIGOSYLTREHALOSE TREHALOHYDROLASE)
DE (TREZ) (EC 3.2.1.141).
GN TREZ.
OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
OX NCBI_TaxID=2287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35092 / DSM 1617 / P2;
RX MEDLINE=21332296; Pubmed=11427726;
RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
RA Awayez M.J., Chan-Weiher C.C.-Y., Clausen I.G., Curtis B.A.,
RA De Moors A., Erauso G., Fletcher C., Gordon P.M.K.,
RA Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
RA Garret R.A., Ragan M.A., Sensen C.W., Van der Oost J.;
RA "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
DR EMBL; Y08256; CAA69503.1;
DR EMBL; AE006815; AAK42272.1;
DR InterPro; IPR000461; Alpha_amylase.
DR InterPro; IPR004193; Isoamylase_N.
DR Pfam; PF00128; alpha-amylase; 1.
DR Pfam; PF02922; isoamylase_N; 1.
KW Hydrolase; Glycosidase; Complete proteome.
SQ SEQUENCE 561 AA; 64370 MW; B00EA03020F6B242 CRC64;

Query Match 82.8%; Score 2467; DB 17; Length 561;
Best Local Similarity 79.7%; Pred. No. 3e-148;
Matches 447; Conservative 63; Mismatches 49; Indels 2; Gaps 1;

OY 1 MTFAYKIDGNEVIFTLWAPYQKSVKLVLEKGLYEMERDEKGYFTITLNNVKVDRYKYV 60
Db 1 MTFAYKIDGNEVIFTLWAPYQKSVKLVLEKGLYEMERDEKGYFTITLNNVKVDRYKYV 60
```

Db 1 MFGYKLEDDGVFNLPWYQKVKLILNRGIYEMERDDKGYFTITLDNRVVGDRYKI 60
QY 61 LDDASEIPDPASRYQPGVHGSPSIIQESKEFNNEF--LKKEDELIYETHVGTFTPEGT 118
Db 61 LDDNSEVPDPASRYQPGVHGSEIISDPFENDSENSYKVKREDLVYIELHIGFTFSEGT 120
QY 119 FEGVIRKLDYKDLGITAIEIMPTIAQPPGKRDGMDGYVLYAVQNSYGGPGRKFLVDEA 178
Db 121 FEGVIRKLDYKDLGITAIEIMPTIAQPPGKRDGMDGYVLYAVQNSYGGPGRKFLVNEA 180
QY 179 HKKGLGLVLDVYVNHVGPENYVVKLGYPYSQKYKTPWGLTFNFDDEAESEVVRKFLVNV 238
Db 181 HKGLGLVLDVYVNHVGPENYVVKLGYPYSQKYKTPWGLTFNFDDEAESEVVRKFLVNV 240
QY 239 EYWIKEYNVDFRLDAVHALIDTSPKHILEEIAADVHKYNRIVIAESDLNDPRVNPKEK 298
Db 241 EYWIKEYNVDFRLDAVHALIDTSPKHILEEIAADVHKYNRIVIAESDLNDPRVNPKEK 300
QY 299 CGYNIDAQWDDFHHSTHAYLTGEROGYTDGFLNDDIVKSKYDVFYDGYKYSNFRKTH 358
Db 301 CGYNIDAQWDDFHHSTHAYLTGEROGYTDGFLNDDIVKSKYDVFYDGYKYSNFRKTH 360
QY 359 GEPVGLDGCNFVYVYQNHDDQVNRGKGERIILKLVDRSKYIAAALYLSPIYIPMFGE 418
Db 361 GKSVDLDDGCKFVYVYQNHDDQVNRGKGERIILKLVDRSKYIAAALYLSPIYIPMFGE 420
QY 419 EYGEENPFYFSDFSKLIQGVREGKKGODTQPDSTFNASKLSWKIDEEIFSFY 478
Db 421 EYGEENPFYFSDFSKLIQGVREGKKGODTQPDSTFNASKLSWKIDEEIFSFY 480
QY 479 KILIKMRKELSIACDRRVNVNGENMLIIKRGREYFSLYVFSKSYEGYSGTLSSNNS 538
Db 481 KSLIKRKEYGLACNRKLSVNGENYVLTGKNGCLAVYVFSKSYEGYSGTLSSNNS 540
QY 539 FPOHIEGKGFYFSDKGFALYKL 559
Db 541 FPOHIEGKGFYFSDKGFALYKL 561

RESULT 4
Q53641
ID Q53641 PRELIMINARY; PRT; 556 AA.
AC Q53641; 008279; 008064;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE ALPHA-AMYLASE (FRAGMENT).
GN TREZ.
OS Sulfolobus acidocaldarius.
OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
OX NCBI_TaxID=2285;
RN [1]
RP SEQUENCE FROM N.A.
RA Kato M., Kettoku M., Miura Y., Komeda T., Konishi Y., Shindo K.,
RL Kobayashi K., Iwanatsu A.;
RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC33909;
RX MEDLINE=97135071; PubMed=8980629;
RA Maruta K., Mitsuzumi H., Nakada T., Kubota M., Chaen H., Fukuda S.,
RT "Cloning and sequencing of a cluster of genes encoding novel enzymes
RT of trehalose biosynthesis from thermophilic archaeobacterium Sulfolobus
RT acidocaldarius.";
RL Biochim. Biophys. Acta 1291:177-181(1996).
DR EMBL; D64131; BAAL1011.1;
DR EMBL; D83245; BAAL1863.1;
DR InterPro; IPR000461; Alpha.amylase.
DR InterPro; IPR0004193; isoamylase.N.
DR InterPro; IPR003682; sub.transporter.
DR Pfam; PF00128; alpha-amylase; 1.
DR Pfam; PF02922; isoamylase_N; 1.

DR PROSITE; PS00217; SUGAR_TRANSPORT_2; UNKNOWN_1.
KW Hydrolase.
FT NON_TER 556 556
SQ SEQUENCE 556 AA; 64373 MW; 4215B45E6C8E4E7 CRC64;

Query Match 58.7%; Score 1748.5; DB 1; Length 556;
Best Local Similarity 58.9%; Pred. No. 8.9e-103;
Matches 329; Conservative 87; Mismatches 136; Indels 7; Gaps 5;

QY 2 TFAYKIDGNEVITLWAPYQKSVKLVKLEKLYEMERDEKGYTITLNNVVDRTKYVL 61
Db 3 SFGNIEKNGKIFKLPWYVNSVKL--LSKKLIPMEKNDEGFEDETEENLTYSII 61
QY 62 DDASEIPDPASRYQPGVHGSPSIIQESKEFNNEF--LKKEDELIYETHVGTFTPEGT 121
Db 62 EDKREIPDPASRYQPGVHGSPSIIQESKEFNNEF--LKKEDELIYETHVGTFTPEGT 121
QY 122 VIRKLDYKDLGITAIEIMPTIAQPPGKRDGMDGYVLYAVQNSYGGPGRKFLVDEA 181
Db 122 VIEKLDYKDLGITAIEIMPTIAQPPGKRDGMDGYVLYAVQNSYGGPGRKFLVDEA 181
QY 182 GLGVLDVYVNHVGPENYVVKLGYPYSQKYKTPWGLTFNFDDEAESEVVRKFLV 241
Db 182 GIADVLDVYVNHVGPENYVVKLGYPYSQKYKTPWGLTFNFDDEAESEVVRKFLV 241
QY 242 IKENYVNDGFRDLDAVHALIDTSPKHILEEIAADVHKYNRIVIAESDLNDPRVNPKEK 301
Db 242 FKTFKIDGLDLDAVHALIDTSPKHILEEIAADVHKYNRIVIAESDLNDPRVNPKEK 301
QY 302 NIDAQWDDFHHSTHAYLTGEROGYTDGFLNDDIVKSKYDVFYDGYKYSNFRKTH 361
Db 300 KIDAQWDDFHHSTHAYLTGEROGYTDGFLNDDIVKSKYDVFYDGYKYSNFRKTH 361
QY 362 VGELDGCNFVYVYQNHDDQVNRGKGERIILKLVDRSKYIAAALYLSPIYIPMFGE 421
Db 360 VGLDPPRKFFVYVYQNHDDQVNRGKGERIILKLVDRSKYIAAALYLSPIYIPMFGE 421
QY 422 EENPFYFSDFSKLIQGVREGKKGODTQPDSTFNASKLSWKIDEEIFSFY 481
Db 420 ETNPFYFSDFSKLIQGVREGKKGODTQPDSTFNASKLSWKIDEEIFSFY 481
QY 482 IKMRKELSIACDRRVNVNGENMLIIKRGREYFSLYVFSKSYEGYSGTLSSNNS 541
Db 480 INIRKRYN--NCKRVKVRREGNCTILMEKIGIIASFDDIVINSKITGNLLIGI--GFPK 536
QY 542 HIEGK-YEFDKGFALYKL 559
Db 537 KLKDELIVKNGVGYQL 555

RESULT 5
Q973H4
ID Q973H4 PRELIMINARY; PRT; 359 AA.
AC Q973H4;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE PUTATIVE ALPHA-AMYLASE.
GN ST0927.
OS Sulfolobus tokodaii.
OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
OX NCBI_TaxID=111955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JCM 10545 / 7;
RX PubMed=11572479;
RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA Sekine M., Baba S.-I., Ankai A., Kosugi H., Hosoyama A., Fukui S.,
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
RA Yoshizawa T., Tanaka T., Kudo Y., Yamazaki J., Kishida N., Oguchi A.,
RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
RA Oshima T., Kikuchi H.;

RT "Complete genome sequence of an aerobic thermoacidophilic
RL Crenarchaeon, Sulfolobus tokodaii strain 7.";
DNA Res. 8:123-140(2001).
DR EMBL; AP000984; BAB5939.1; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 359 AA; 41600 MW; 4DF429099EB1665 CRC64;

Query Match 42.6%; Score 1269; DB 17; Length 359;
Best Local Similarity 64.9%; Pred. No. 1.1e-72;
Matches 233; Conservative 51; Mismatches 71; Indels 4; Gaps 3;

Qy 201 MVKLGPTFSQYKYPWGLTFNFDNDAESDEVKRFLENYVEYKYNVDGFLDAVHAIIID 260
Db 1 MFWLGPVFSQYKYPWGLTFNFDNDAESDEVKRFLENYVEYKYNVDGFLDAVHAIIID 60
Qy 261 TSPKHLEIAADVHKYRIVIAESDLNDPRVNPKECGKYNIDAQWDDPHHSIAYLT 320
Db 61 SSPKHILQDIAELSHSGKGFVIAESDLNDPKIID- -DKCGYKIDAQWDDPHHSVHAYLT 118
Qy 321 GERQYITDFGNLDIIVKSYKDFVYDCKYSNFRKTHGEPVGLDGCNFFVYIQNHQDV 380
Db 119 GERNYSYDFGLDIIIVKAFKDFVYDCKYSNFRKTHGAPVGNLACKFVYIQNHQDV 178
Qy 381 GNRKGERIKLVRESYKIAAALYLLSPYIPMIFMGEEYGEENPFYFSDSKLIQ 440
Db 179 GNRGDRSLVLVQKSYMIASALYLLSPYIPMIFMGEEYGEENPFYFSDSKLIQ 238
Qy 441 VREGRKENGODTPODESTFNASKLWKIDEEIFSVFKILMKRKLSTACDRRVNVN 500
Db 239 VREGRKNGOSIDPQSDAEFLKSLKSNIDKEMLEYRTLIKVKFNGKCSRVEVDK 298
Qy 501 GENLIIRKGYFSLYFVSSEIVKSYSGTLLSNNSFPQIHIEGK-YEPDKGFALYK 558
Db 299 TDNLVIGLEKIFLLVAFSDGLILRYNGELFSTSD-FPKRIYSSKEIKVKGVYK 356

RESULT 6

Q9AD15 PRELIMINARY; PRT; 581 AA.
ID Q9AD15
AC Q9AD15
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-OCT-2001 (Tremblrel. 18, Last annotation update)
DE PUTATIVE ALPHA AMYLASE.
GN SCBAC1A6.02C.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Seeger K.J., Harris D.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Cardeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RX MEDLINE=97000351; PubMed=8843436;
RA Redenbach M., Kieser H.M., Denapaita D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
DR EMBL; AL589708; CAC33923.1; -
DR InterPro; IPR000461; Alpha_amylase.
DR Pfam; PF00128; alpha-amylase; 1
SQ SEQUENCE 581 AA; 63952 MW; 8616DF3BC1A3408 CRC64;

Query Match 34.7%; Score 1033.5; DB 2; Length 581;
Best Local Similarity 40.3%; Pred. No. 1.7e-57;
Matches 213; Conservative 92; Mismatches 185; Indels 39; Gaps 13;
Qy 14 FTLWAPYQSKVKLVKLEKLYEMERDEK--GYFTITLNNVKRD--RYKYVLDASEIPD 69
Db 3 FEVWAPQAGRVTLR-CDGATRALERDPERPGW--CCEAPARDGSRGYFAVDGPPVLPD 58
Qy 70 PASRYQEGVHGPGSQIIQESKEFNNEFLKKEDL--IIYEHVGTTPTECTFEGVIRKL 126
Db 59 PRSRQDGPGLDSAVVDHGR-YAWTPWAGRPLPGGVLYELHVGTYTPEGLDAAADR 117
Qy 127 DYLDKDLGTTATEIMPIAQFPGKRDWGDVYLYAVQNSYGGPEGRKLVDAHKKGLGVI 186
Db 118 EHLVRLGVTHVELMPLCFPGRGHGWEGYSLWAVEHPYGGPEALKRFVDRAHDLGLGV 177
Qy 187 LDVYVNHVGPENYMKVLPYFSQYKYPWGLTFNFDNDAESDEVKRFLENYVEYKYN 246
Db 178 LDVYVNHLPSPGNHLPAGFPYFTDHTTPWGVAVNLDPAGSDEVRAVYLDVSALAWLRDY 237
Qy 247 VDGRLDAVHAIIIDTSKPHILEEIAVYH-----KYNRIVIAESDLNDPRVNPKECG 300
Db 238 LDGRLDAVHALADTRACHFEQLSTAVDGLAADLERPLFLIAESDLNDPRIITPRAEGG 297
Qy 301 YNIDAQWDDPHHSIAYLTGERQYITDFGN--LDDIVKSYKDFVYDCKYSNFRKTH 358
Db 298 LGVHAQWDDPHHHAALTGESQGYADFARDPLAALAKTLIRGYFDGTYSFGRSH 357
Qy 359 GEPV--GELGCGNFVYIQNHQVGNRGKGERIKLVRESYKIAAALYLLSPYIPMIF 416
Db 358 GRALDRGVAHRLTGYSTQHDQVGNRAQGRDLASVSPGLAACATLTLTAPTPLMFM 417
Qy 417 GEEYGEENPFYFSDSKLIQVREGKKE-----NGODT-DPODESTFNASKLSWK 469
Db 418 GEWAAGTQWFFDHTDPLQAQAVROGRREFAAHGWREEDVDPDQDPATRERSCLDWS 477
Qy 470 IDE-----EIFSYFKILIKRKEL-----SIACDRRVNVNNGENWLIK 508
Db 478 EPEREHARVLWDYRRLIALRAEQPLTDPDLATRVAYDGTGWLAFR 526

RESULT 7

Q9RX51 PRELIMINARY; PRT; 600 AA.
ID Q9RX51
AC Q9RX51
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE MALTOOLIGOSYLTREHALOSE TREHALOHYDROLASE, PUTATIVE.
GN DR0464.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R1;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "genome sequence of the radioresistant bacterium Deinococcus
radiodurans R1.";
RL Science 286:1571-1577(1999).
DR EMBL; AE001905; AAF10042.1; -
DR TIGR; DR0464; -
DR InterPro; IPR000461; Alpha_amylase.
DR InterPro; IPR004193; isoamylase_N.

DR Pfam; PF00128; alpha-amylase; 1.
DR Pfam; PF02922; isoamylase N; 1.
KW Hydrolase; Complete proteome.
SQ SEQUENCE 600 AA; 66909 MW; 594091EC093F8A44 CRC64;

Query Match 30.4%; Score 905.5; DB 16; Length 600;
Best Local Similarity 38.7%; Pred. No. 2.3e-49;
Matches .205; Conservative 86; Mismatches 170; Indels 69; Gaps 20;

QY 14 FTLWAPYQSKVKLV-----LEKGLYMERDEKGYFTIILNNVYVDRKYKYLDDA 64
DB 38 FLWTSTARTAVRVNGTEHVMTSLGGIYELEP-----VGPARYLFVLDGV 86
QY 65 SELPDASRYQPGVGPSPQIIQESKEFNNEFR-----LKKEDLIIEYHVGTFP 115
DB 87 -PTPDYARELPDGVGEAEV---DFG--TFDWTADWHGKIKLADCFYEVHVGTFP 139
QY 116 EGTFFGVIRKLDYKLDGITAIEIMPIAOPGKRDNGYGVLYAVONSYGSGPEGRKLV 175
DB 140 EGYTAAAEKLPYLKELGVTAIQVMPLAAPDQGRGWYDGAAYAPYAPYGRPEDLMALV 199
QY 176 DEAHKKGVLGVLDVYVNHVCPGNYVKLGP-YFSOKYKTPWGLTFNFDDAESDEVKFI 234
DB 200 DAAHRLGLGVLDVYVNHVCPGNYLSSYAPSYFTDFRSSAWGM--GLDYAE-PHMRVY 256
QY 235 LENVEYWIKEYNVDGRFLDAVHAIIDTSPKHILEEIAVVHXY--NRIVIAESDLNDPRV 292
DB 257 TGNARWLRYDFDGLRLDTPYTDSDSETHTLTAQELHELGGTHLLAEDHRLPDL 316
QY 293 VNPKECGYNDIAQWDDFHSHYLTGERQGYTDF-GNLDIDVYKVDVYDQKYS 351
DB 317 VTWN-----HLGDITWDDFHETRVTLTGEGYAGYRGGAALAYTTRRGWRYEGQFW 371
QY 352 NFRKTH--GEPVGLDGCNFVYVYIONHDQVGNRGKGERIKL--VDRESYKIAAALYLL 407
DB 372 AVGEEHEHGHPSDALEAPNFYVICIIONHQIGNRJLGERLHOSDGVTLHEYRGAAALL- 430
QY 408 SPYIPIMFGEYGEENPFYFFSDFSKLIQVREGKRKE-----NGQDT-DPQDES 459
DB 431 -PMTLLFOGEWAASTPQFFSDHA-GELQAVSSEGRKKEFGSGFSGEDVPDPAEQ 488
QY 460 TFNASKLSKWIDE-----EIFSYKILIKMKELSTACDR-RVNVNGEN 503
DB 489 TFLNSKLINLAERGGHARTLRYDLRLRLRRDPVHLNRQENRLTGH 538

RESULT 8
Q911V1 PRELIMINARY; PRT; 583 AA.
AC Q911V1;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-OCT-2001 (Tremblrel. 18, Last annotation update)
DE PROBABLE GLYCOSYL HYDROLASE.
GN PA2164.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Saiter M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen."
RA Nature 406:959-964(2000).

DR EMBL; AE004643; AAG05552.1; --
DR InterPro; IPR002086; Aldehyde_dehydr.
DR InterPro; IPR000461; Alpha_amylase.
DR Pfam; PF00128; alpha-amylase; 1.
DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
KW Hydrolase; Complete proteome.
SQ SEQUENCE 583 AA; 65657 MW; 651CD7A918833E00 CRC64;

Query Match 29.6%; Score 881.5; DB 16; Length 583;
Best Local Similarity 36.3%; Pred. No. 7.3e-48;
Matches 203; Conservative 95; Mismatches 174; Indels 87; Gaps 19;

QY 3 FAYKIDGN-EVIFTLWAPYQSKVKLVKELGYMERDEKGYFTIILNNVYVDRKYKYL 61
DB 7 FGAQFOGNGRTCEGLWAPDAREVRVETADGRWPLSGSDSGWFEATL-PCPCTRYRYRI 65
QY 62 DDASETPDASRYQPGVGPSPQIIQES-----KEFNNEFLKKEDLIIEYHVGTFP 115
DB 66 DGRPGVPDPASQFPDGVHGSQVLDHGTYAVRWDEWRGPW---HEAVIYELHVGLE-- 120
QY 116 EGTFFGVIRKLDYKLDGITAIEIMPIAOPGKRDNGYGVLYAVONSYGSGPEGRKLV 175
DB 121 -GSYAEVERPLRLVELGVTAVELMPLGEPGRRNMGYDGVLPFAPASAYGTPEQLKHLI 179
QY 176 DEAHKKGVLGVLDVYVNHVCPGNYMVK-LGPFYFSOKYKTPWGLTFNFDDAESDEVKFI 234
DB 180 DSAHGGLAMVFDVYVNHVCPGNYLAQYAAAFRRDROTGWQAIDF---RRGEVREFF 236
QY 235 LENVEYWIKEYNVDGRFLDAVHAIIDTSPKHILEEIA-----DVVHYKNRIVIAE 284
DB 237 YENALMWLLDYRVDFLDAVHAIPDSA---FLVEMARLRGAAGPERHVH-----LVLE 288
QY 285 SLDLNDPRVNPKECGYNDIAQWDDFHSHYLTGERQGYTDFGN-LDDIVKSYKDV 343
DB 289 NDNRRASLL---RQGY--DAQMNDGHHALVLLTGENDGYQDYDPEPLRCLARCLAG 342
QY 344 FVYDGYSNFRRTKTHGEPVGLDGCNFVYVYIONHDQVGNRGKGERIKLVDRSEYKIAA 403
DB 343 FVYQGE-ANRHGRPCEPSADLAPDAFVFLQNHQDQVGNRAFRGERSLVLAEPQALRLA 401
QY 404 LYLLSYIPIMFGEYGEENPFYFFSDFSKLIQVREGKRKEGQD-----452
DB 402 LQLLAMPILPLFMGECAAREPFLYFTD-HQGLADAVREGRKKEFGFEGEGATLAS 460
QY 453 -TDPQDESTFNASK-----LSWKIDEEIFSYKILIKMKELST-----490
DB 461 LPDPNAVETFRSPGLADCDPAWR-----GFYROLLEIRHEHLIPRLRGARSLGVTTI 514

QY 491 ---ACDRRVNVVNGENWLI 506
DB 515 AGAALSARWRLGDSGSDRI 533
RESULT 9
Q92U63 PRELIMINARY; PRT; 601 AA.
AC Q92U63;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE PUTATIVE 1,4-ALPHA-GLUCAN BRANCHING ENZYME PROTEIN (EC 2.4.1.18).
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Plasmid pSymb (megaplasmid 2).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396508; PubMed=11481431;
RA Finan T.M., Weidner S., Wong K., Buhrmester J., Chain P.,

Db 64 EDOVDVTPYARDIDES-SGKDNSIARIKDGKEIKVDTYVWQHDDKPLPADHELVIYEL 122
QY 109 HVGFET-----PEGFEGWIRKLDYLGKLTGTAIEMPTAOPGKRDGCVGLYAV 161
Db 123 HVGDFSGGDDPYARGKYKHVIEKLDYLCGLINAIELLPVKEYPGDYSWGNPRYEFAT 182
QY 162 QNSYGGPEGRFKLVDEAHKGLGLVDVYNHVGPEGNY-MVKLGPYFSQKYKTP---WG 217
Db 183 ESSVGSATDLKLVDECHQIRIIMDGIYNHSEASSPLTQIDHDYWHHHEPRDPDNWG 242
QY 218 LTFNF-----DDAESDEVKRFILENVWEIKYVNDGFRDAVHAIDTSPKHIL-----EE 269
Db 243 PEFNYHYDENLETYPARKFGTGVRYWVGEYHLDGIRYDAARQIANFYDFMHWIAQBAKK 302
QY 270 IADVVHKNRIVTAESDLNDPRVNV---PKEKCGYNIDAQWDDFHSHIAYLTGERQGY 326
Db 303 TAGAKPYN---VAEHIPETTSITNLGPMDCG-----WHDSEFYHTIKAHICGDT--- 349
QY 327 YTOFGNLDIDVKSVDYVVDGKYSNFRKTHGEPVGLDGCNPFVVIQNHG-----QV 380
Db 350 -FDLENLKDVI-----DPKROGF-----LGATNVVNYLTNHDHIMVEL 388
QY 381 GNRGKGRIIKLVDRSY-----KIAAALYLLSPYIPMTFMGEYGEENPFYFSDSKL 437
Db 389 GNR-----EIRHDEAFRAKLGTAIIMTAVGVPLIWMGEFGEYKRP----- 429
QY 438 IOGVREGKKGODTDPQDESTFNASKLSWKIDEEFYSFKILIKMRK 486
Db 430 -----KQDDQSKIDWTLLGNDLNRSLFDYHKGLIGLRK 462

RESULT 12
Q59242

ID Q59242 PRELIMINARY; PRT; 652 AA.
AC Q59242: 008485;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE 1.4-ALPHA-GLUCAN BRANCHING ENZYME (EC 2.4.1.18)
DE (GLYCOPEN BRANCHING ENZYME) (1.4-ALPHA-GLUCAN BRANCHING ENZYME)
DE (AMYLO-(1,4 TO 1,6)TRANSGLUCOSIDASE)
DE (AMYLO-(1,4-1,6)-TRANSGLUCOSYLASE).
GN GLGB.
OS Bacillus stearothermophilus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Geobacillus.
OX NCBI_TaxID=1422;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=TRBE14;
RC MEDLINE=95031021; PubMed=7944355;
RX Takata H., Takaha T., Kuriki T., Okada S., Takagi M., Imanaka T.;
RA "Properties and active center of the thermostable branching enzyme
RT from Bacillus stearothermophilus.";
RL Appl. Environ. Microbiol. 60:3096-3104(1994).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=TRBE14;
RA Takata H.;
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RN SEQUENCE FROM N.A.
RC STRAIN=TRBE14;
RA Takata H., Takaha T., Okada S., Takagi M., Imanaka T.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: FORMATION OF 1,6-GLUCOSIDIC LINKAGES OF
CC GLYCOPEN.
CC -1- PATHWAY: THIRD STEP IN GLYCOPEN BIOSYNTHESIS.
CC -1- MISCELLANEOUS: OPTIMAL ACTIVITY AT APPROXIMATELY 39 DEGREES
CC CELSIUS.
CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.

DR EMBL; D87026; BAA19588.1; -.
DR InterPro; IPR000461; Alpha-amylase.
DR InterPro; IPR004193; Isoamylase.N.
DR Pfam; PF00128; alpha-amylase; 1.
DR Pfam; PF02922; isoamylase.N; 1.
KW Glycogen biosynthesis; Transference; Glycosyltransferase.
FT ACT_SITE 308 308 BY SIMILARITY.
FT ACT_SITE 351 351 BY SIMILARITY.
FT ACT_SITE 419 419 BY SIMILARITY.
SQ SEQUENCE 652 AA; 76793 MW; 4591EB414A0E3FEF CRC64;

Query Match 13.4%; Score 399; DB 2; Length 652;

Best Local Similarity 25.2%; Pred. No. 3e-17;

Matches 149; Conservative 96; Mismatches 189; Indels 158; Gaps 31;

QY 4 AYKIDGNEVI-----FTLWAPYQKSVKL-----KVLEKGLYEMER-DEKGYFTITLN 49
Db 22 SYELFGAHVINEGKVGTRFCVWAPHAREVRLVSGFNDWDGTDRLKVKDEGVWTVVVP 81

QY 50 NVKVRDRYKY--VLDDASEI--PDASRYQPEGVHGSPQIIQESK--EFNNEFLKK--- 100
Db 82 ENLEGHLYKYEIVTPDQGVLFKADPTAFYSELRPHTAS-IAYDLKGQWMDQSWKKRRR 140

QY 101 -----EDLIIEIHVGTF-TPEGTF-----EGVIRKLDYLDKLGITATEIMPIAQFPGKRD 150
Db 141 KRIYDQPMVIYELHFGSKKKDKGRFYTYREMADELISYVLDHGFTHELLPLVEHPDLRS 200

QY 151 WGTGCVLYAVQNSYGGPEGRKLVDEAHKGLGLVDVYNHVV--GPEGNYVMVKLGP--- 206
Db 201 WGYQGTGYAVTSRYGTPHDFMYFVDRCHQAGIGVIMDWVPGHFCKDAHGLYMFDCGPTY 260

QY 207 -YFSOKYKT--PWGLTFNFDDAESDEVKRFILENVWEIKYVNDGFRDAVHAII--- 259
Db 261 EYANEKDRENYVWG-TANF-DLGKPEVRSFLISNALFWLEYIHDGFRVDAVANMLYWP 318

QY 260 -----DTSPKHILEIADVVHVKY--NRIVIAESDLNDPRVWNPKEKCGYNIDAQWVDDF 311
Db 319 NDRLYENPYAVEFLRKLNEAVFAYDPNALMAEDSTDWPKVTAPTYEGGLGFYKWK--- 374

QY 312 HHSIAYLTGEROYTYDFGNLDDIVKSYKDFVYDGKYFNFRKTHGEPVGLDGCNFFV 371
Db 375 -----NMGMNDMLK-YMETPPYE-----RRHVHNVQVTFSL---LY 406

QY 372 VYIQN-----HDQVGNRGGERIILVDRESYKIAAALYLLSPYI-----PMIEMGE 419
Db 407 AYSENFILPFSHDEVVH-GKKSLLNKMPGSYEERF-AQLRLLYGYMMAHPGKLLFMGNE 464

QY 420 YGEENPFYFFSDFSLSKLIQGVREGKKGKENGQDTPQDESTFNASKLSWKIDBEI----- 474
Db 465 -----FAQFDE-----WKFEDLDWVLF 482

QY 475 -FSFYKILIKMRKELSTACDRRVNVV-----NGENWLLIKREYFSLVVF 518
Db 483 DFLHRKMNDYMKEL-IACYKRYKPFYELDHPDQGFWDHVAHQEQ-SIFS 532

RESULT 13
Q93Q35

ID Q93Q35 PRELIMINARY; PRT; 440 AA.
AC Q93Q35;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE BRANCHING ENZYME GLGB (FRAGMENT).
OS Myxococcus xanthus.
OC Bacteria; Proteobacteria; delta subdivision; Myxobacteria;
OC Myxococcales; Cystobacterineae; Myxococcaceae; Myxococcus.
OX NCBI_TaxID=34;
RN [1]
RN SEQUENCE FROM N.A.
RA Ueki T., Inouye S.;
RT "Identification of a new His-Asp phosphorelay signal transduction

RT system which regulates expression of a heat shock gene, lonD, of
RT Myxococcus xanthus.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF285783; AAK83002.1; -;
FT NON_TER 440 440
SQ SEQUENCE 440 AA; 48964 MW; 8D492EA5A4A92017 CRC64;

Query Match 13.3%; Score 397; DB 2; Length 440;

Best Local Similarity 28.0%; Pred. No. 2.3e-17;

Matches 131; Conservative 75; Mismatches 164; Indels 98; Gaps 23;

```
QY 14 FTLMAPYQKSVLKVLEKLY-----EMERDEKGYFTITLNNVVRDRYKYL----- 61
Db 12 FRWAP-----MASRVFSGDFNGWTWELGNEFNFGSGDVAGAVKQGYKFTIRNQWG 67
QY 62 DDAEIPDPASRYQEGVHGSPQIIQESKEF--NNETFLKK--EDLIIEIHVGTFT--TP 115
Db 68 SDAWR-ADPRSQAQENST--GSSIIYDHGEYWNAAQQYSSPGCFNEMIIYEMHVGTFHDS 124
QY 116 ---EGTFEGVIRKLDLGLGTAIEIMPIAOPGKRWDGVDVLYAVQNSYGGPEGFR 172
Db 125 GFGPGNWSAIALRDLRGANMIKMPAYEFAGDFSGWYNAAPFAPESAYGHPNDMK 184
QY 173 KLVDEAHKKGLGVILDVYNNHVP-----EGNYMVKLGPYFSQYK--TPWGLTFNF 222
Db 185 REVDEAHMRGIGVIFDVVHNNHGPSDLPMWCFSGDCLSGGGEYFYNDWRKSTPWGDT--R 242
QY 223 DDAEIDEVRKFTLENVYKWEYNNVDGFRDLAVHAI-----IDTSPKHILEEIAVDVHKY 277
Db 243 PDYGRPEVRAIRDSMMNLTHSFRDGLRWDTKYMTQNGSDTT---AIPDAMRVFRSI 299
QY 278 NR-----IVIAESDLNDPRVYN---PREKCGYNIDAQWDDFHSHIAYLTGERQG 325
Db 300 NREINATQPKISIAEDFGGDFITNDATSDTSGCAGFSQWGGDFVHAIRAAVIAS--- 356
QY 326 YTTDFGNLDDIVKSKVDFV--YDKYSNFRKRTGEPVGLDGCNFVYIQNHQDVGNR 383
Db 357 --NDSGR---DNMSVRNATORISGRHT-----ARVYSESHDEVAN- 393
QY 384 KGRERIKLV-----DREYKIAAALYLLSPYIPMIFMGERYGE 422
Db 394 GKA-RVPEINPGNAGSNAKKRSTLAAGVVTSPGIMFIQGGEFLE 440
```

RESULT 14

```
Q930Z9
ID Q930Z9 PRELIMINARY; PRT; 1142 AA.
AC Q930Z9;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ALKALINE PULLULANASE.
GN PUL1876.
OS Bacillus sp. KSM-1876.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=72413;
(1)
RN SEQUENCE FROM N.A.
RC STRAIN=KSM-1876;
RA Hatada Y., Ico S.;
RT "alkaline pullulanase from the alkaliphilic bacterium.";
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB049812; BAB47586.1; -;
SQ SEQUENCE 1142 AA; 128759 MW; 318FC01362961E84 CRC64;
```

Query Match 13.3%; Score 395.5; DB 2; Length 1142;

Best Local Similarity 22.9%; Pred. No. 1.1e-16;

Matches 164; Conservative 100; Mismatches 190; Indels 261; Gaps 34;

QY 8 DGNEVITFLWAP-----YQR-----SVKLVLEKGLYEMERDE-----KG 42

```
Db 376 DGTATI-KLWSPKAEHVQVILYDKNDPDDIVTEVEMKLGDRGVWEVQLTKEKNTGLDSLGR 434
QY 43 YFTITLNNVVRDRYKYLVD--DASEIPDPASR-----YQPEGVHGSPQIIQ 87
Db 435 YV-----YHYEIDHDGDKKIALDPYAKSLSAWSNDEQGPYAKAALVDPSSIGP 482
QY 88 ESKEFNNETFLKKEDLIIEIHVGTFTPE-----GTTFEGVIRKLDYKLDGITA 136
Db 483 ELEFAHIEGFEKKEDLIIEVHVDRFTSDPHIADeltaTQAGTFASFVKLDYIEDLGYTH 542
QY 137 IEIMPIAOP-----PKRD-----WGVDGYLYAVQNSYG---GPE---G 170
Db 543 IQLPVMVSFFANEFKNDERMLDFESTNTNMGYDPSQYFALTGMYSIEDTDPDLRIKE 602
QY 171 FRKLVDEAHKKGLGVILDVYNNHVPPEGNYMVKLGPYFSQYKTPWGLTFNFDDAES--- 227
Db 603 PKKLIDELHSRGMGVVLVDVYNNHRTARVIGIFEDLVNY-----YHFMADMGTPR 650
QY 228 -----DEVKFFILENVYKWEYNNVDGFRDLAVHAIIDTSPKHILEEIAVDVH 275
Db 651 TSFGGRLGTTHEMSRRILVDISITVWVEEFKIDGFRPDM-----GDHDAESIQAID 703
QY 276 KY-----NRVIAE-----SDLNDPRVYNPKKCGYNIDAQWV-----DDFHHS 314
Db 704 KAKEINPNVIMIGEWIITAGDEDDPNV-----QAADQHWMOYTESVGVFSDEFERNE 755
QY 315 IHA-----YLTGERQGYTDFGNLD--DIVKSKVDFVYDKYSNFRKRTGTHGPV 362
Db 756 LKSGFGEHGEPRFTGGAR-----NIDLFDNKAQPHNFIAD----- 793
QY 363 GELDCNPFVYIQNHQDVGNRGKGERIILKVDV-----ESYKIAAALYLLSPYIP 412
Db 794 --DPGDVVPYIEAHD---NLTLHDVITAMAIQKDDHDEEIHQRIRLGNLMTLSOGIA 847
QY 413 MIFMGEEY-----EEN--PF---YFFSDFSSKLGQVREGKK 447
Db 848 FLHAGQYGRTKQFRAETSEPPYKSTYMTDENGEPFRYPYFIHDSYDSTDIINFDWERA 907
QY 448 ENGODTPODEST-----FNASKLSWK--IDEEIFSYFKILIKMRKELSTAC 492
Db 908 TNA-DAPYQNLTRYTTGLIELRRSSDAFRLGTHKDLVDEKV---TQNLNPEIETDLV 963
QY 493 DRRVNVNGENWLIIKREYFSLVYFSKSSIEVYKSTGLLSSNNSPQHIEEGK 547
Db 964 AYRIEATG-----AFYVFNADDEER---TLTEEDLTVGFEFVVDCK 1004
```

RESULT 15

```
Q59319
ID Q59319 PRELIMINARY; PRT; 825 AA.
AC Q59319;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ALPHA-DEXTRIN 6-GLUCANOHYDROLASE (EC 3.2.1.41).
GN PUL1A.
OS Caldcellum saccharolyticum (Caldicellulosiruptor saccharolyticus).
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Thermoaerobacter group; Caldicellulosiruptor.
OX NCBI_TaxID=44001;
(1)
RN SEQUENCE FROM N.A.
RA Albertson G.D., McHale R., Gibbs M.D., Bergquist P.L.;
RT "Cloning and expression of a type II pullulanase from an extremely
thermophilic anaerobic bacterium, Caldicellulosiruptor
saccharolyticus.";
RL Eur. J. Biochem. 0:0-0(1995).
(2)
RN SEQUENCE FROM N.A.
RA Gibbs M.D.;
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; L39876; AAB06264.1; -;
```

Search completed: July 15, 2002, 12:24:08
Job time: 390 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 12:17:32 ; Search time 58.93 Seconds
(without alignments)
1047.973 Million cell updates/sec

Title: US-09-298-924-8
Perfect score: 2954
Sequence: 1 MFSGGNIENKGIKFLWAP.....KLKKDELIKVNRGVGVQLE 556

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 11073796 residues
Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802.*
1: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES							
Result	Query						
No.	Match	Score	Length	ID		Description	
1	100.0	2954	556	17	AA92755	Trehalose-releasin	
2	100.0	2954	556	17	AA90620	Sulfolobus acidoca	
3	59.2	1748.5	559	17	AA90619	Sulfolobus solfata	
4	35.3	1042.5	595	22	AA92072	C glutamicum prote	
5	35.3	1042.5	610	22	AA80206	Corynebacterium gl	
6	32.8	969	589	16	AA80289	Trehalose releasin	
7	32.8	969	596	16	AA77470	Trehalose releasin	
8	32.5	959	598	16	AA77471	Trehalose releasin	
9	32.4	958	597	16	AA80290	Trehalose releasin	
10	31.7	936	575	21	AA85157	Trehalose-releasin	
11	31.7	936	575	21	AA85165	Trehalose-releasin	

12	360	12.2	652	16	AA80037	Bacillus stearothe
13	360	12.2	652	17	AA96109	Starch-branching-e
14	356	12.1	658	13	AA923787	Heat-resistant pul
15	342.5	11.6	639	12	AA911271	B.stearothermophil
16	336	11.4	731	22	AA911100	C glutamicum prote
17	336	11.4	731	22	AA979423	Corynebacterium gl
18	326.5	11.1	648	10	AA94635	B. thuringiensis a
19	323	10.9	772	18	AA94567	Thermotoga maritim
20	321	10.8	772	19	AA949871	Thermotoga maritim
21	320	10.8	793	20	AA900869	S. tuberosum isoam
22	313.5	10.6	1250	21	AA91279	Group B Streptococ
23	312	10.6	670	22	AA60913	Propionibacterium
24	310.5	10.5	606	18	AAW36602	S. tuberosum debra
25	306	10.4	718	22	AAU60903	Propionibacterium
26	306	10.4	766	20	AAU00870	S. tuberosum isoam
27	303.5	10.3	666	20	AA934991	C. pneumoniae prot
28	301	10.2	562	10	AA91904	Sequence of amy B
29	298.5	10.1	621	21	AA91285	A polypeptide with
30	296.5	10.0	818	20	AA917523	Zea mays SU1 starc
31	296	10.0	630	22	AA69074	Aquifex aeolicus V
32	293.5	9.9	931	20	AA927357	Group B Streptococ
33	292	9.9	829	22	AAE05691	Bacillus acidopull
34	292	9.9	862	21	AA978513	Pullulanase (pulB)
35	285	9.6	764	21	AA950819	Wheat isoamylase p
36	284	9.6	798	20	AAW73552	Pullulanase protei
37	284	9.6	1938	20	AAW73553	Full length Pullul
38	282	9.5	893	17	AAW09257	Bacillus alkaline
39	282	9.5	1938	17	AAW09255	Bacillus alkaline
40	280.5	9.5	931	20	AA900872	Original S. tubero
41	278.5	9.4	618	20	AA937184	Protein involved i
42	277	9.4	921	22	AAE05723	Bacillus acidopull
43	277	9.4	921	22	AAE05726	Bacillus acidopull
44	276	9.3	762	21	AAW09077	N. denitrificans a
45	276	9.3	921	22	AAE05689	Bacillus acidopull

ALIGNMENTS

RESULT 1.
AA92755
ID AA92755 standard; Protein; 556 AA.
AC
XX AA92755;
XX
XX
DT 03-AUG-1996 (first entry)
XX
XX
DE Trehalose-releasing thermostable enzyme.
XX
KW Thermostable enzyme; trehalose; sweetener.
XX
XX
OS Sulfolobus acidocaldarius strain ATCC 33909.
XX
PN AU9527131-A.
XX
XX
PD 01-FEB-1996.
XX
PF 21-JUL-1995; 95AU-0027131.
XX
XX
PR 04-JUL-1995; 95JP-0189760.
PR 21-JUL-1994; 94JP-0190180.
PR 11-APR-1995; 95JP-0109128.
XX
XX (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
XX
XX
PI Kubota M, Mitsuzumi H, Sugimoto T;
XX
XX
DR WPI; 1996-106284/12.
DR N-PSDB; AAT16899.
XX
XX
PT Recombinant thermostable enzyme from Sulfolobus acidocaldarius, -
PT releases trehalose from non-reducing saccharide at temps. exceeding
PT 55 degrees Centigrade

```

XX PS Claim 2; Page 53-54; 74pp; English.
XX PS
XX CC A thermostable enzyme (AAR92755) of Sulfolobus acidocaldarius ATCC
CC 33909 releases trehalose from non-reducing saccharides having a
CC trehalose structure as an end unit and a degree of polymerisation of
CC at least 3. It has a mol.wt. of 54-64 kDa (SDS-PAGE), a pI of
CC 5.6-6.6 and is substantially not inactivated when incubated in aq.
CC solution (pH 7.0) at 85 deg for 60 min. Recombinant enzyme is
CC obtd. by expression of an isolated DNA fragment (AAT16899) in
CC host cells, pref. Escherichia coli, using e.g. vector pBluescript
CC II SK(+). The trehalose is useful as a sweetener.
XX SQ Sequence 556 AA;

Query Match 100.0%; Score 2954; DB 17; Length 556;
Best Local Similarity 100.0%; Pred. NO. 7.5e-267;
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MFSFGGNIERNKGIFKLWAPYVNSVKLSKLLIPMEKNDEGFEVEIDDIENLTYSYI 60
Db 1 mfsfggniekngifklwapyvnsvklskllipmekndegffevelddeenltysyi 60

Qy 61 IEDKREIPDPASRYOPLGVHDKSQLIRTDYQILDGLGVKIEDLIYELHVGTFSGEGNFK 120
Db 1 iedkreipdpasryoplglvhdksqlirtdyqildglgvkiedliiyelhvgtfsgegnfk 120

Qy 121 GVIEKLDYLKDLGTIGTIELMPVAQFPNGNRDYGVDGLYAVQNTYGGPWELAKLVNEAHK 180
Db 121 gviekldylkdlgtigtietlmpvaqfpngnrwydgvlavqntyygppwelaklvneahk 180

Qy 181 RGIAVILDVYVNHIGPEGNYLLGLSPYFSDRYKTPWGLTFNFDGDCDQVRKFILENVEY 240
Db 181 rgiauildvvyvnhigpegnyllglspfyfsdryktpwgltnfnfdgdcqvrvkfilenvey 240

Qy 241 WFKTFKIDGLRLDAVHAIFDQVGNRGNGERLSILTDTKTYLMAATLYILSPYILFMGEYYE 420
Db 241 wfktdkldglrldavhaifdnspkhilqetaekahqlgkfviasesdlnpkivkddcgyk 420

Qy 421 TNPFFFSDFSDPVLIKVREGRLKENNQMDPOSEAFKLSKLSWKIDEEVLDDYKQLI 480
Db 421 tnpfffsdfsdpvlkivregrlkennqmdpqseeflksklswkideevlddykqli 480

Qy 481 NIKRYNNCKRVKVRREGNCITLIMEKIGTIIASFDIVINSKITGNLLIGIGFPKLIK 540
Db 481 nirkrynnckrvkvrregncitlimekigtiiasfddivinskitgnlligigfppklikk 540

Qy 541 DELIKVRNGVGYQLE 556
Db 541 delikvrngvgyqle 556

RESULT 2
AAR90620
ID AAR90620 standard; Protein; 556 AA.
XX AC
XX AC AAR90620;
XX DT
XX DT 29-JUN-1996 (first entry)
XX XX Sulfolobus acidocaldarius amylase for alpha, alpha-trehalose prodn.
XX DE transferase: amylase; Sulfolobus; production; alpha, alpha-trehalose;
XX KW malto-oligosaccharide; hydrolysis.
KW
```

```

XX OS Sulfolobus acidocaldarius.
XX OS
XX PN WO9534642-A.
XX PN
XX PD 21-DEC-1995.
XX PD
XX PF 14-JUN-1995; 95WO-JF01189.
XX PF
XX PR 21-APR-1995; 95JP-0120673.
XX PR
XX PR 15-JUN-1994; 94JP-0133354.
XX PR
XX PR 18-AUG-1994; 94JP-0194223.
XX PR
XX PR 31-OCT-1994; 94JP-0290394.
XX PR
XX PR 21-NOV-1994; 94JP-0286917.
XX PR
XX PR 21-NOV-1994; 94JP-0311185.
XX PR
XX PA (KIRI ) KIRIN BEER KK.
XX PA
XX DR WPI; 1996-049671/05.
XX DR
XX DR N-PSDB; AAT12326.
XX DR
XX PT Sulfolobus spp. derived transferase and amylase - for production of
XX PT alpha, alpha-trehalose from malto-oligosaccharide(s)
XX PT
XX PS Claim 112; Page 244-250; 357pp; Japanese.
XX PS
XX CC The amylase is derived from Sulfolobus acidocaldarius. The amylase acts
CC on a saccharide having at least three sugar units, which are pref.
CC glucose units at the reducing end (the linkage between the first and
CC second glucose units is alpha-1, alpha-1, while the linkage between the
CC second and third glucose units is alpha-1,4), to hydrolyse alpha-1,4
CC linkages within the sugar chain, yielding alpha, alpha-trehalose and also
CC mono- and disaccharide hydrolysis products. The amylase has a mol. wt. of
CC 61 to 64 kDa. It is characterised by working at pH 4.5-5.5 and at 60-85
CC deg.C. It has an isoelectric point of 4.3-5.4 and retains at least
CC 100 percent activity after 6 hrs. at 80 deg.C. It is completely inhibited
CC by 5 mM copper sulphate. Use of a transferase and the amylase in
CC succession on suitable substrates such as malto-oligosaccharides, is
CC useful for the production of alpha, alpha-trehalose.
XX CC
XX SQ Sequence 556 AA;

Query Match 100.0%; Score 2954; DB 17; Length 556;
Best Local Similarity 100.0%; Pred. NO. 7.5e-267;
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MFSFGGNIERNKGIFKLWAPYVNSVKLSKLLIPMEKNDEGFEVEIDDIENLTYSYI 60
Db 1 mfsfggniekngifklwapyvnsvklskllipmekndegffevelddeenltysyi 60

Qy 61 IEDKREIPDPASRYOPLGVHDKSQLIRTDYQILDGLGVKIEDLIYELHVGTFSGEGNFK 120
Db 61 iedkreipdpasryoplglvhdksqlirtdyqildglgvkiedliiyelhvgtfsgegnfk 120

Qy 121 GVIEKLDYLKDLGTIGTIELMPVAQFPNGNRDYGVDGLYAVQNTYGGPWELAKLVNEAHK 180
Db 121 gviekldylkdlgtigtietlmpvaqfpngnrwydgvlavqntyygppwelaklvneahk 180

Qy 181 RGIAVILDVYVNHIGPEGNYLLGLSPYFSDRYKTPWGLTFNFDGDCDQVRKFILENVEY 240
Db 181 rgiauildvvyvnhigpegnyllglspfyfsdryktpwgltnfnfdgdcqvrvkfilenvey 240

Qy 241 WFKTFKIDGLRLDAVHAIFDQVGNRGNGERLSILTDTKTYLMAATLYILSPYILFMGEYYE 420
Db 241 wfktdkldglrldavhaifdnspkhilqetaekahqlgkfviasesdlnpkivkddcgyk 420

Qy 421 TNPFFFSDFSDPVLIKVREGRLKENNQMDPOSEAFKLSKLSWKIDEEVLDDYKQLI 480
Db 421 tnpfffsdfsdpvlkivregrlkennqmdpqseeflksklswkideevlddykqli 480

Qy 481 NIKRYNNCKRVKVRREGNCITLIMEKIGTIIASFDIVINSKITGNLLIGIGFPKLIK 540
Db 481 nirkrynnckrvkvrregncitlimekigtiiasfddivinskitgnlligigfppklikk 540

Qy 541 DELIKVRNGVGYQLE 556
Db 541 delikvrngvgyqle 556
```


Db 361 gdlpprkfvvfnqhdqvgngngersiltdkttlylmaatlylspylplfmgeeyye 420

QY 421 TNPFFFFSDPVLKIGVREGRLKENNOMIDPQSEAFKLSKLSWKIDEEVLDYKQLI 480

Db 421 tnpfffsdvpvlkigvregrlkennqmidpqqseafklsklskwideevldyvkli 480

QY 481 NIKRYNNCKRVKEVRREGNCITLIMEKIGITIASFDIVINSKITGNLLIGIFPKKLKK 540

Db 481 nlrkrynnckrvkevrregncitlimekigilasfddivinskitgnlligifpkklkk 540

QY 541 DELIKVNRGVGVYQLE 556

Db 541 delikvnrvgvgyqle 556

RESULT 3

AAR90619

ID AAR90619 standard; Protein; 559 AA.

AC AAR90619;

XX

DT 29-JUN-1996 (first entry)

XX

DE Sulfolobus solfataricus amylase for alpha, alpha-trehalose prodn.

XX

KW transferase; amylase; Sulfolobus; production; alpha, alpha-trehalose;

KW malto-oligosaccharide; hydrolysis.

XX

OS Sulfolobus solfataricus.

XX

PN W09534642-A.

XX

PD 21-DEC-1995.

XX

PF 14-JUN-1995; 95WO-JP01189.

XX

PR 21-APR-1995; 95JP-0120673.

PR 15-JUN-1994; 94JP-0133354.

PR 18-AUG-1994; 94JP-0194223.

PR 31-OCT-1994; 94JP-0290394.

PR 21-NOV-1994; 94JP-0286917.

PR 21-NOV-1994; 94JP-0311185.

XX

PA (KIRI) KIRIN BEER KK.

XX

DR WPI; 1996-049671/05.

DR N-PSDB; AAT12325.

XX

PT Sulfolobus spp. derived transferase and amylase - for production of

PT alpha, alpha-trehalose from malto-oligosaccharide(s)

XX

PS Claim 108; Page 235-240; 357pp; Japanese.

XX

CC The amylase is derived from Sulfolobus solfataricus. The amylase acts

CC on a saccharide having at least three sugar units, which are pref.

CC glucose units at the reducing end (the linkage between the first and

CC second glucose units is alpha-1, alpha-1, while the linkage between the

CC second and third glucose units is alpha-1,4), to hydrolyse alpha-1,4

CC linkages within the sugar chain, yielding alpha, alpha-trehalose and also

CC mono- and disaccharide hydrolysis products. The amylase has a mol. wt. of

CC 61 to 64 kDa. It is characterised by working at pH 4.5-5.5 and at 60-85

CC deg.C. It has an isoelectric point of 4.3-5.4 and retains at least

CC 100 percent activity after 6 hrs. at 80 deg.C. It is completely inhibited

CC by 5 mM copper sulphate. Use of a transferase and the amylase in

CC succession on suitable substrates such as malto-oligosaccharides, is

CC useful for the production of alpha, alpha-trehalose.

XX

SQ Sequence 559 AA;

Query Match 59.2%; Score 1748.5; DB 17; Length 559;

Best Local Similarity 58.9%; Pred. No. 2.9e-154;

Matches 329; Conservative 87; Mismatches 136; Indels 7; Gaps 5;

QY 3 SFGGNIEKNKGFIFKLWAPVNSVYKLGK-LSKKLIPMEKNDGFEFEIDDEENLTYSYII 61

Db 2 tfaykidgnevifltlwapyqskvklvlekglyemerdekgyftitlnnvkvdrdrykyl 61

QY 62 EDKREIPDPASRYOPIGVHDKSOLIRTDVQILDGKVKIEDLLIYELHVGTFSSQGNFKG 121

Db 62 ddaseipdpasryqpegvghpsqlideskefnnetfikkedlliyehvgtftpegtfeg 121

QY 122 VIEKLDYLDKLTGTIELMPVAQPGNRDMGYDGVFLYAVONTYGGPWELAKLVNBAHAKR 181

Db 122 virkldyldkltgtielmpiaqfpgkrdwgydgvlyavqnsyggpegfrkrlvdeahkk 181

QY 182 GIAVILDVYVNHIGPEGNYLLGLGYPFSDRYKTPWGLTFNFDRCGDQVAKFILENVEYW 241

Db 182 glgvildvynhvpegnyvmvklgpyfsqyktpwgltfnfdaesdevrkfilenveyw 241

QY 242 FKTFKIDGLRLDAVHAIFDNPKHILQETAERAKAHLGKFKVIAESDLNDPKIV--KDDCGY 299

Db 242 ikeynvdgfrldavhaidtspkhileetadvvhkynrlviaesdlndprvnpkckcy 301

QY 300 KIDAQWDDFHHAVHAFITKEKDYIYQDFGRIEDIEKTKDFVYDGYKSYRGRTHGAP 359

Db 302 nidaqwwddfhhsihayltgerggytdfgnlddiwksykdvfvydgkysnfrkrthgep 361

QY 360 VGDLPFRKFVVFIONHDQVGNRNGRERLSILTDKTYLMAATLYLISPYILFIMGEYY 419

Db 362 vgelgcnfvvyiqnhdqvgngrgkeriiklvdesykaaalyllsypilmfmggeyg 421

QY 420 ETNPFFFFSDPVLKIGVREGRLKENNOMIDPQSEAEFLKLSWKIDEEVLDYKQL 479

Db 422 eenpfyffsdfsdkliqvgvregkrkngqtdpqdestfnasklswkideeifsfykl 481

QY 480 INIRKRYN-NCKRVKEVRREGNCITLIMEKIGITIASFDIVINSKITGNLLIGI--GFPK 536

Db 482 ikmrkelsiactdrvnvngengenwiikgreysiyvfkssievkysgtllsnnnsfpq 541

QY 537 KLKXDELKIKVNRGVGVYQL 555

Db 542 hieegk-yefdkgfalykl 559

RESULT 4

AAG92072

ID AAG92072 standard; Protein; 595 AA.

AC AAG92072;

XX

DT 26-SEP-2001 (first entry)

XX

DE C glutamicum protein fragment SEQ ID NO: 5826.

XX

KW Corynebacterium protein; amino acid synthesis; vitamin; saccharide;

KW organic acid synthesis.

XX

OS Corynebacterium glutamicum.

PN EP1108790-A2.

XX

PD 20-JUN-2001.

XX

PF 18-DEC-2000; 2000EP-0127688.

XX

PR 16-DEC-1999; 99JP-0377484.

PR 07-APR-2000; 2000JP-0159162.

PR 03-AUG-2000; 2000JP-0280988.

XX

PA (KYOW) KYOWA HAKKO KOGYO KK.

XX

PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;

PI Tateishi N, Senoh A, Ikeda M, Ozaki A;

XX

DR WPI; 2001-376931/40.

DR N-PSDB; AAH67291.
XX Novel polynucleotides derived from Corynebacterium bacteria, for identifying
PT mutation point of a gene, measuring expression of a gene, analysing
PT expression profile or pattern of a gene and identifying homologous gene
XX
PS Claim 17: SEQ ID NO: 5826; 246pp + Sequence Listing; English.
XX
XX The present invention provides a number of nucleotide and protein
CC sequences from the Corynebacterium bacterium Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of corynebacterium bacterium, measuring expression amount and
CC analysing the expression profile or expressing pattern of a gene derived
CC from Corynebacterium bacterium, and identifying a homologue of a gene derived
CC from corynebacterium bacterium. Corynebacterium bacteria are useful for producing
CC amino acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a protein described
CC in the exemplification of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC European Patent Office.
XX
SQ Sequence 595 AA;

Query Match 35.3%; Score 1042.5; DB 22; Length 595;
Best Local Similarity 40.8%; Pred. No. 2.7e-88;
Matches 225; Conservative 93; Mismatches 199; Indels 35; Gaps 13;
Qy 15 FKLWAPYVNSVKLSKLLIPMEKNDGFEVEIDDIENITYSYIID- - - - -KREIPDP 70
Db 28 fswwaplphdvhlngtelpmhktgswraeap-kagdrygfslfdgsswsktlpdp 86
Qy 71 ASRVQPLGVHDKSLQIRTDYQILD- - - - -LGKVKIEDLIYELHVGTFSGEGNFKGVIEKLD 127
Db 87 rstsqpdgvglhsevsddsywgdqwtgtri-lpgsvlyelhvgtfsgdtfegvvdklp 145
Qy 128 YKLDLGTIGIELMPVAOPPPGNRDCWYGVLYAVQNTYGGPWEIAKLVAEAKRGIAVIL 187
Db 146 ylrldgvtailpvpqfpgnrgwydgvlwhavhagvggpaglklldashqagiavyl 205
Qy 188 DVVYNHGPEGNYLLGLGPYFSDRYKTPWGLTFNFDRCQDQVRKFTILENVEYWFKTFKI 247
Db 206 dvvynhfgpdyngyqfpytsqg-stgwgdvvnghdsdevrnylldaarqwfedfhv 264
Qy 248 DGLRLDAVHAIFDNPSPKHLQIEAKAHLQ- - - - -KFVIAESDLNDPKIV- - - - -KDDCGY 299
Db 265 dglrlдавhsldrgaysllaqltmvaedvsagtgprsliaeseldpdkftvsreaggf 324
Qy 300 KIDAQWVDDFHVAFITKEDKYYQDFGRIDIEKTFKDFVYVDGKYSRYGRTHGAP 359
Db 325 gldaqvddihhalvsgernyysdfgsvdtlaktlrevfhtgnytygrnhrp 384
Qy 360 VGD- - -LPPRKVFVFIQNHQVGNRNGERLS-ILTDKTTYLMAATLYILSPYIPLFMGE 416
Db 385 vhpditpasrfvtyttthdqtgnraigdrpstltlpeqgikaaaliy-sspytgmfnge 443
Qy 417 EYETNPFFSDFSDPVLKIGVREGRLKE- - - - -NNQMIDPQSEAFKLSLWKID 469
Db 444 efgattpfafcshtdpelntitsegkrrefarigwnaddipspelestftskldwft 503
Qy 470 EE- - -VLDYYKOLINIRKRYNNCK- - - - -RVKVRREGNCTITLMEKIGIASFDDIVINSK 523
Db 504 aeqrindayqllhlrhtlqfsgpnlltlevehgenwlsmangrgillanfsddtitvp 563
Qy 524 ITGNLLIGIGPP 535
Db 564 lggeliysftsp 575

RESULT 5
AAB80206

ID AAB80206 standard; Protein; 610 AA.
XX
AC AAB80206;
XX
DT 30-APR-2001 (first entry)
XX
DE Corynebacterium glutamicum MP protein sequence SEQ ID NO:1146.
XX
KW Corynebacterium glutamicum; metabolic pathway protein; MP protein;
KW fine chemical production; microorganism; organic acid; nucleoside;
KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleotide;
KW lipid; saturated fatty acid; unsaturated fatty acid; diol; vitamin;
KW carbohydrate; aromatic compound; cofactor; polyketide; enzyme.
XX
OS Corynebacterium glutamicum.
XX
PN WO200100843-A2.
XX
PD 04-JAN-2001.
XX
PF 23-JUN-2000; 2000WO-IB00923.
XX
PR 25-JUN-1999; 99US-0141031.
PR 01-JUL-1999; 99DE-1030476.
PR 02-JUL-1999; 99US-0142101.
PR 08-JUL-1999; 99DE-1031415.
PR 08-JUL-1999; 99DE-1031418.
PR 08-JUL-1999; 99DE-1031419.
PR 08-JUL-1999; 99DE-1031420.
PR 08-JUL-1999; 99DE-1031424.
PR 08-JUL-1999; 99DE-1031428.
PR 08-JUL-1999; 99DE-1031434.
PR 08-JUL-1999; 99DE-1031435.
PR 08-JUL-1999; 99DE-1031443.
PR 08-JUL-1999; 99DE-1031453.
PR 08-JUL-1999; 99DE-1031457.
PR 08-JUL-1999; 99DE-1031465.
PR 08-JUL-1999; 99DE-1031478.
PR 08-JUL-1999; 99DE-1031510.
PR 08-JUL-1999; 99DE-1031541.
PR 08-JUL-1999; 99DE-1031573.
PR 08-JUL-1999; 99DE-1031592.
PR 08-JUL-1999; 99DE-1031632.
PR 08-JUL-1999; 99DE-1031634.
PR 08-JUL-1999; 99DE-1031636.
PR 08-JUL-1999; 99DE-1032125.
PR 08-JUL-1999; 99DE-1032126.
PR 08-JUL-1999; 99DE-1032130.
PR 08-JUL-1999; 99DE-1032186.
PR 08-JUL-1999; 99DE-1032206.
PR 08-JUL-1999; 99DE-1032227.
PR 08-JUL-1999; 99DE-1032228.
PR 08-JUL-1999; 99DE-1032229.
PR 08-JUL-1999; 99DE-1032230.
PR 14-JUL-1999; 99DE-1032922.
PR 14-JUL-1999; 99DE-1032926.
PR 14-JUL-1999; 99DE-1032928.
PR 14-JUL-1999; 99DE-1033004.
PR 14-JUL-1999; 99DE-1033005.
PR 14-JUL-1999; 99DE-1033006.
PR 12-AUG-1999; 99US-0148613.
PR 27-AUG-1999; 99DE-1040764.
PR 27-AUG-1999; 99DE-1040765.
PR 27-AUG-1999; 99DE-1040766.
PR 27-AUG-1999; 99DE-1040832.
PR 31-AUG-1999; 99DE-1041378.
PR 31-AUG-1999; 99DE-1041379.
PR 31-AUG-1999; 99DE-1041380.
PR 31-AUG-1999; 99DE-1041394.
PR 31-AUG-1999; 99DE-1041396.
PR 03-SEP-1999; 99DE-1042076.
PR 03-SEP-1999; 99DE-1042077.
PR 03-SEP-1999; 99DE-1042079.

PR 03-SEP-1999; 99DE-1042086.
PR 03-SEP-1999; 99DE-1042087.
PR 03-SEP-1999; 99DE-1042088.
PR 03-SEP-1999; 99DE-1042095.
PR 03-SEP-1999; 99DE-1042124.
PR 03-SEP-1999; 99DE-1042129.
PR 09-MAR-2000; 2000US-0187970.
XX (BADI) BASF AG.
XX Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;
XX WPI; 2001-137957/14.
DR N-PSDB; AAF72325.
XX Nucleic acids from Corynebacterium glutamicum encoding metabolic
PT pathway proteins, useful for producing fine chemicals in
PT microorganisms, including organic acids, nonproteinogenic amino acids,
PT and purine and pyrimidine bases -
XX Claim 20; Page 1722-1725; 1737pp; English.
XX AAF71753 to AAF72330 encode the Corynebacterium glutamicum metabolic
CC pathway (MP) proteins given in AAB79634 to AAB80211: The C. glutamicum
CC MP nucleic acids are useful for the production of fine chemicals
CC in microorganisms, including organic acids, nonproteinogenic amino
CC acids, purine and pyrimidine bases, nucleosides, nucleotides, lipids,
CC saturated and unsaturated fatty acids, diols, carbohydrates, aromatic
CC compounds, vitamins, cofactors, polyketides and enzymes.
XX Sequence 610 AA;
SQ
Query Match 35.3%; Score 1042.5; DB 22; Length 610;
Best Local Similarity 40.8%; Pred. No. 2.8e-88;
Matches 225; Conservative 93; Mismatches 199; Indels 35; Gaps 13;
QY 15 FKLWAPYVNSVKLKSLLKIPMEKDEGFVEIDIEENLTYSYIIE---KREIPDP 70
DB 43 fswaplphdvhilngeltpmhktgswraeiap-kagdrygfslfdgsswsktlpdp 101
QY 71 ASRYQLGVHDKSOLRTDYQILD---LGKVKIEDLIIYELHVGTFSGNFKGVIEKLD 127
DB 102 rstqpgdvghlsvsdssdysdsgdqwtgri-lpgsvlyelhvgtfsgdftgfvvdkip 160
QY 128 YLKDLGITGLMPLVAFQFNGRDMGVDGVFLYAVQNTYGGPWLAKLVNEAKRGIAVIL 187
DB 161 Yirdlgvtaiellpvqfggnrnwgdvhlwhavhagyg9pgagkklldashqaglavyl 220
QY 188 DVVYNHIGPEGNYLLGLGYFSDRYKTPMGLTFNFDRCDDQVRKFIENVEYFFTKFI 247
DB 221 dvvynhfgpgnynqgfpysg-stgwgdvvnvngdhdsdevrnyildaarqwfedfhv 279
QY 248 DGLRLDAVHAIFDPSKPHILOETAEAKHQLG-----KVFIAESDLNDPKIV---KDCGY 299
DB 280 dgirldavhslldrgayslaqltmvaedvsagtgiprsliaeslndpkfvtsreaggf 339
QY 300 KIDAQWDDFHAVHAFITKEKDYDYODFGRIEDIEKTFKDVYDGYKSYRGRGTHGAP 359
DB 340 gidaqvwddihhahlvsgerngyysdfgsdvtlaktirevftehtgnytygrnhgtp 399
QY 360 VGD--LPPRKVFQIHNQVQNGRNGERLS-ILTDKTYLMAATYILSPYIPLIFMGE 416
DB 400 vhpdiptasarfvtthdqtgnraigdrpstltltpceqqvkaaiiy-spytpcmifmge 458
QY 417 EYETNPFFFSDFSPVLKGVREGRLE-----NNQIMDPQSEEAFLSKLSWKID 469
DB 459 efgattpaffcshtdpelnrltsegrkrefarlgwnaddipselstftssklgweft 518
QY 470 EE---VLDYKQLINIRKRYNCK---RVKEVRREGNCITLIMEKIGIITASFDIVNSK 523
DB 519 aeqrindaykqlilhrltglfsqpnlltlevehgenwismangrilansfddtitvp 578

QY 524 ITGNLLIGIGFP 535
DB 579 Iggellysftsp 590
RESULT 6
ID AAR80289 standard; Protein; 589 AA.
XX AAR80289;
XX 19-JAN-1996 (first entry)
XX Trehalose releasing enzyme.
KW Enzyme; trehalose; alpha-glucosyltrehalose; alpha-maltosyltrehalose;
KW alpha-maltotriosyltrehalose; alpha-maltotetraosyltrehalose;
KW maltopentaosyltrehalose; sweeter; taste-improving agent;
KW stabiliser; filler; excipient; adjuvant; foodstuffs; cosmetics;
KW pharmaceuticals.
XX Rhizobium sp. M11.
XX EP671470-A2.
XX 13-SEP-1995.
XX 07-MAR-1995; 95EP-0301474.
XX 07-MAR-1994; 94JP-0059840.
XX 07-MAR-1994; 94JP-0059834.
XX (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
XX Hattori K, Kubota M, Sugimoto T, Tsusaki K;
XX WPI; 1995-312772/41.
XX N-PSDB; AAQ98669.
PT DNA encoding a trehalose releasing enzyme - which releases trehalose
PT from a non-reducing saccharide having a trehalose structure as an
PT end unit.
XX Claim 3; Page 21-22; 45pp; English.
XX This enzyme can be used for the preparation of trehalose with high
CC yields and efficiency from non-reducing saccharides such as
CC alpha-glucosyltrehalose, alpha-maltosyltrehalose,
CC alpha-maltotriosyltrehalose, alpha-maltotetraosyltrehalose and
CC maltopentaosyltrehalose. The trehalose can be used as a sweetener,
CC taste-improving agent, quality-improving agent, stabiliser, filler,
CC excipient or adjuvant in food products cosmetics and pharmaceuticals.
XX Sequence 589 AA;
SQ
Query Match 32.8%; Score 969; DB 16; Length 589;
Best Local Similarity 39.5%; Pred. No. 2e-81;
Matches 215; Conservative 74; Mismatches 195; Indels 60; Gaps 12;
QY 13 GIFKLWAPYVNSVKLKSLLKIPMEKN-----DEGFVEIDIEENLTYSYIIE-DKR 65
DB 8 grfdwaapeagtvtllageryemgrrpgnpgadegwtaadaptaadvgyllgdgel 67
QY 66 ETPDPSRQPLGVHDKSOLRTDYQILDGLKVKIEDL-----IYELHVGTFSQ 115
DB 68 plpdprrtrrqqegvh---alsrt-----fdpgahrwdagwqgrelgsgviyelhigtft 120
QY 116 EGNFKGVIEKLDYLDLKITGIELMPVAFQFNGRDMGVDGVFLYAVQNTYGGPWLAKLV 175
DB 121 egtldaaagklidyaglgldfclpvnafngthnwgvgdvgwfvahngygg9paaygrfv 180
QY 176 NEAKRGIAVILDVYNHIGPEGNYLLGLGYFSDRYKTPMGLTFNFDRCDDQVRKFI 235

Db 181 daahaaglgvldvvyvnhlgsgnylprygyllkhgegnwgdsvnlpgpsdhvrgyil 240
 Qy 236 ENVEYFTEKIDGLRLDAVHAIFDNPSPKHILQETIAEKAHOLGK-----FVIAESDLND 289
 Db 241 dnvamwlrdivrvgdlridavhalkderavhileefgaladalsesegrpptliaesdlnd 300
 Qy 290 PKIV--KDDCGYKIDAQVDDFHHAHAFITKEKDYQQDFGRIDIEKTFKOVFVYDVK 347
 Db 301 prllprdvngyglagqsdffhahvhnvsggettyysdfdsigalakvldgdfhdgs 360
 Qy 348 YSRVGRTHGAPV--GDLPPRKFFVFTQNDQVNCNRCNGERLSLTLDKTYLMAATLYIL 405
 Db 361 yssfrgrchrpfnfsavhpaaalvvcsgnhdqignratgdrslsgslpygsalaaavltlt 420
 Qy 406 SPYIPLFMGEYETNPFFFSDFSDPVLKGVREGRLKENNOM-----IDPQSEEA 458
 Db 421 gpftpmfmggeygattpwqfftshepeelgkataegrirefermgwdpavvpqdpqpet 480
 Qy 459 FLKSLSWK-----IDEEVDYKQLINIRKRYNNCKRVKREVRREGNCITLIMEKIGIIA 513
 Db 481 ftrskldwaeasagdhallllyrsltlr-----fstpelarlglfadtav----- 526
 Qy 514 SFDD 517
 Db 527 efdd 530

RESULT 7
 AAR77470
 ID AAR77470 standard; Protein; 596 AA.
 XX
 AC AAR77470;
 XX
 DT 25-JAN-1996 (first entry)
 XX
 DE Trehalose releasing enzyme.
 XX
 KW Enzyme; trehalose; alpha-glucosyltrehalose; alpha-maltosyltrehalose;
 KW alpha-maltotriosyltrehalose; alpha-maltotetraosyltrehalose;
 KW maltopentaosyltrehalose; sweetener; taste-improving agent;
 KW stabiliser; filler; excipient; adjuvant; foodstuffs; cosmetics;
 KW pharmaceuticals.
 XX
 OS Rhizobium sp. M11.
 XX
 PN EP671470-A2.
 XX
 PD 13-SEP-1995.
 XX
 PF 07-MAR-1995; 95EP-0301474.
 XX
 PR 07-MAR-1994; 94JP-0059840.
 PR 07-MAR-1994; 94JP-0059834.
 XX
 PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
 XX
 PI Hattori K, Kubota M, Sugimoto T, Tsusaki K;
 XX
 DR WPI: 1995-312772/41.
 DR N-PSDB; AAQ98671.
 XX

XX DNA encoding a trehalose releasing enzyme - which releases trehalose
 PT from a non-reducing saccharide having a trehalose structure as an
 PT end unit.
 XX
 PS Claim 3; Page 27-29; 45pp; English.
 XX

XX This enzyme can be used for the preparation of trehalose with high
 CC yields and efficiency from non-reducing saccharides such as
 CC alpha-glucosyltrehalose, alpha-maltosyltrehalose,
 CC alpha-maltotriosyltrehalose, alpha-maltotetraosyltrehalose and
 CC maltopentaosyltrehalose. The trehalose can be used as a sweetener,
 CC

CC taste-improving agent, quality-improving agent, stabiliser, filler,
 CC excipient or adjuvant in food products cosmetics and pharmaceuticals.
 XX
 SQ Sequence 596 AA;

Query Match 32.8%; Score 969; DB 16; Length 596;
 Best Local Similarity 39.5%; Pred. No. 2e-81;
 Matches 215; Conservative 74; Mismatches 195; Indels 60; Gaps 12;

Qy 13 GIFKLWAPVNSVVKLSKLLPMEKN-----DEGFEVEIDIEENLTYSYLIE-DKR 65
 Db 15 grfdwapeagvtvllaggyemgrrpgngpadegwtaadaptdadvdygylldgdei 74
 Qy 66 EIDPPASRYOPLGVHDKSQLIRTDYQILDGLGVKIEDL-----IYELHWGTFQSO 115
 Db 75 plpdrtrrpegvh---alsrt-----fdpgahrwqdagwggrelgsvlyelhigftfp 127
 Qy 116 EGNFKGVIEKLDYLDKDLGITGIELMPVAQFPNGRDWGYGVFLYAVQNTYGGPWELAKLV 175
 Db 128 egtldaaagklidyaglgidfiellpvnafngthnwygvgwfvahveggygpaaygrfv 187
 Qy 176 NEAHKRGIAVILDVYVNHIGPEGNYLLGLGPYFSDRYKTPMGLTFNFDRCDDQVRKFI 235
 Db 188 daahaaglgvldvvyvnhlgsgnylprygyllkhgegnwgdsvnlpgpsdhvrgyil 247
 Qy 236 ENVEYFTEKIDGLRLDAVHAIFDNPSPKHILQETIAEKAHOLGK-----FVIAESDLND 289
 Db 248 dnvamwlrdivrvgdlridavhalkderavhileefgaladalsesegrpptliaesdlnd 307
 Qy 290 PKIV--KDDCGYKIDAQVDDFHHAHAFITKEKDYQQDFGRIDIEKTFKOVFVYDVK 347
 Db 308 prllprdvngyglagqsdffhahvhnvsggettyysdfdsigalakvldgdfhdgs 367
 Qy 348 YSRVGRTHGAPV--GDLPPRKFFVFTQNDQVNCNRCNGERLSLTLDKTYLMAATLYIL 405
 Db 368 yssfrgrchrpfnfsavhpaaalvvcsgnhdqignratgdrslsgslpygsalaaavltlt 427
 Qy 406 SPYIPLFMGEYETNPFFFSDFSDPVLKGVREGRLKENNOM-----IDPQSEEA 458
 Db 428 gpftpmfmggeygattpwqfftshepeelgkataegrirefermgwdpavvpqdpqpet 487
 Qy 459 FLKSLSWK-----IDEEVDYKQLINIRKRYNNCKRVKREVRREGNCITLIMEKIGIIA 513
 Db 488 ftrskldwaeasagdhallllyrsltlr-----fstpelarlglfadtav----- 533
 Qy 514 SFDD 517
 Db 534 efdd 537

RESULT 8
 AAR77471
 ID AAR77471 standard; Protein; 598 AA.
 XX
 AC AAR77471;
 XX

DT 25-JAN-1996 (first entry)
 XX
 DE Trehalose releasing enzyme.
 XX

XX Enzyme; trehalose; alpha-glucosyltrehalose; alpha-maltosyltrehalose;
 KW alpha-maltotriosyltrehalose; alpha-maltotetraosyltrehalose;
 KW maltopentaosyltrehalose; sweetener; taste-improving agent;
 KW stabiliser; filler; excipient; adjuvant; foodstuffs; cosmetics;
 KW pharmaceuticals.
 XX

OS Arthrobacter sp. Q36.
 XX

PN EP671470-A2.
 XX

PD 13-SEP-1995.
 XX

```
PF 07-MAR-1995; 95EP-0301474.
XX
XX 07-MAR-1994; 94JP-0059840.
XX 07-MAR-1994; 94JP-0059834.
XX
PA (HAYB ) HAYASHIBARA SEIBUTSU KAGAKU.
XX
PI Hattori K, Kubota M, Sugimoto T, Tsusaki K;
XX
XX WPI; 1995-312772/41.
XX N-PSDB; AAQ98672.
XX
PT DNA encoding a trehalose releasing enzyme - which releases trehalose
PT from a non-reducing saccharide having a trehalose structure as an
PT end unit.
XX
XX Claim 6; Page 27-29; 45pp; English.
XX
XX This enzyme can be used for the preparation of trehalose with high
XX yields and efficiency from non-reducing saccharides such as
XX alpha-glucosyltrehalose, alpha-maltosyltrehalose,
XX alpha-maltotriosyltrehalose, alpha-maltotetrasyltrehalose and
XX maltopentaosyltrehalose. The trehalose can be used as a sweetener,
XX taste-improving agent, quality-improving agent, stabiliser, filler,
XX excipient or adjuvant in food products cosmetics and pharmaceuticals.
XX
XX Sequence 598 AA;
XX
Query Match 32.5%; Score 959; DB 16; Length 598;
Best Local Similarity 40.2%; Pred. No. 1.7e-80;
Matches 202; Conservative 81; Mismatches 186; Indels 34; Gaps 10;
QY 15 FKLWAPYVNSVKLKSLKLIPEK-----NDEGFEVEIDIEENLTVSYIIE-DKREI 67
Db 19 ydwapnaesvtllageryamrraetpedagwwtaagaptgndvgyllldgdetpl 78
QY 68 PDPASRYQLGVHDKSOLI-RTDYQILD---LGKVKIEDLIYELHVGTFSGNFKGYI 123
Db 79 pdptrtrqpdgvhalstfdpsayswgdawqgr-elggavlyelhlgtftpegtleaaa 137
QY 124 EKLDYLKDLGITGIELMPVAQFPGNRDWDGVLVAVQNTYGGPWELAKLVNEAHKRG 183
Db 138 gkldylaglgvdfiellpvnafngthnwgdygvqwfavheayggpeaygrfvdaahaag 197
QY 184 AVILDVYVNHGPEGNYLLGLGPFYSDRYKTPWGLTFNFDRCDOVRKFIENVEYWF 243
Db 198 gviqdvvyhnlpgsgnylprfpgylkqgegnwtgdsvnldgpgsdhvirryildniamwlr 257
QY 244 TFKIDGLRLDAVHAIFDPSFKHILQIEAKHQLGRFV-----IAESDLNDPKIV--KD 295
Db 258 dyrvdglrlidavhalkderavhillefdgaladqisaevgrptliaesdlnnprllyprd 317
QY 296 DCGYKIDAQWDDFHAVHAFITKEKDYIYQDFGRIEDIEKTFKDVVFDGYSRYRGRT 355
Db 318 vngygiegqwsddfhavhvnvtgettygsdfdsalalavlrdffhdgysysfreh 377
QY 356 HGAPV--GDLPPRKFFVFIONHDVGNRNGERLSILTDTKTYLMAATLYILSPYIPLIF 413
Db 378 hgrpinfoavhaalvccsqnhdqignratgdrisqlpygsalalaaviltltpfcpml 437
QY 414 MGEYYETNPFFSDFPVLTKGVREGRLKENNM-----IDPQSEEAFLKSLSW 466
Db 438 mgeeygastpwwftshpepelgkacaeagrikefermgwdpavvpdpqpetfrskldw 497
QY 467 KIDEE-----VLDYYKQLNIRK 484
Db 498 aeaagdhallielysltalr 520
RESULT 9
ID AAR80290 standard; Protein; 597 AA.
```

```
XX
XX AAR80290;
XX
XX 19-JAN-1996 (first entry)
XX
XX Trehalose releasing enzyme.
XX
XX Enzyme; trehalose; alpha-glucosyltrehalose; alpha-maltosyltrehalose;
XX alpha-maltotriosyltrehalose; alpha-maltotetrasyltrehalose;
XX maltopentaosyltrehalose; sweetener; taste-improving agent;
XX stabiliser; filler; excipient; adjuvant; foodstuffs; cosmetics;
XX pharmaceuticals.
XX
XX Arthrobacter sp. Q36.
XX
XX EP671470-A2.
XX
XX 13-SEP-1995.
XX
XX 07-MAR-1995; 95EP-0301474.
XX
XX 07-MAR-1994; 94JP-0059840.
XX 07-MAR-1994; 94JP-0059834.
XX
XX (HAYB ) HAYASHIBARA SEIBUTSU KAGAKU.
XX
XX Hattori K, Kubota M, Sugimoto T, Tsusaki K;
XX
XX WPI; 1995-312772/41.
XX N-PSDB; AAQ98670.
XX
XX DNA encoding a trehalose releasing enzyme - which releases trehalose
XX from a non-reducing saccharide having a trehalose structure as an
XX end unit.
XX
XX Claim 3; Page 24-25; 45pp; English.
XX
XX This enzyme can be used for the preparation of trehalose with high
XX yields and efficiency from non-reducing saccharides such as
XX alpha-glucosyltrehalose, alpha-maltosyltrehalose,
XX alpha-maltotriosyltrehalose, alpha-maltotetrasyltrehalose and
XX maltopentaosyltrehalose. The trehalose can be used as a sweetener,
XX taste-improving agent, quality-improving agent, stabiliser, filler,
XX excipient or adjuvant in food products cosmetics and pharmaceuticals.
XX
XX Sequence 597 AA;
XX
Query Match 32.4%; Score 958; DB 16; Length 597;
Best Local Similarity 40.2%; Pred. No. 2.1e-80;
Matches 202; Conservative 81; Mismatches 186; Indels 34; Gaps 10;
QY 15 FKLWAPYVNSVKLKSLKLIPEK-----NDEGFEVEIDIEENLTVSYIIE-DKREI 67
Db 18 ydwapnaesvtllageryamrraetpedagwwtaagaptgndvgyllldgdetpl 77
QY 68 PDPASRYQLGVHDKSOLI-RTDYQILD---LGKVKIEDLIYELHVGTFSGNFKGYI 123
Db 78 pdptrtrqpdgvhalstfdpsayswgdawqgr-elggavlyelhlgtftpegtleaaa 136
QY 124 EKLDYLKDLGITGIELMPVAQFPGNRDWDGVLVAVQNTYGGPWELAKLVNEAHKRG 183
Db 137 gkldylaglgvdfiellpvnafngthnwgdygvqwfavheayggpeaygrfvdaahaag 196
QY 184 AVILDVYVNHGPEGNYLLGLGPFYSDRYKTPWGLTFNFDRCDOVRKFIENVEYWF 243
Db 197 gviqdvvyhnlpgsgnylprfpgylkqgegnwtgdsvnldgpgsdhvirryildniamwlr 256
QY 244 TFKIDGLRLDAVHAIFDPSFKHILQIEAKHQLGRFV-----IAESDLNDPKIV--KD 295
Db 257 dyrvdglrlidavhalkderavhillefdgaladqisaevgrptliaesdlnnprllyprd 316
QY 296 DCGYKIDAQWDDFHAVHAFITKEKDYIYQDFGRIEDIEKTFKDVVFDGYSRYRGRT 355
```


Db 258 ptveyanekdrenyvwg-tanf-dlgkpevrsflisnalfwleyhldgfrvdavnmly 315
Qy 256 ----HAIFDNS-PKHILQEIAEK--AHOLGKFVIAESDLNDPKIVKD-----DCGYKID 302
Db 316 wpndrilyenpyaveflrkineavfaydpnalmiaedstdwkvptaptgyeglgfnykwn 375
Qy 303 AQWVDDF-----HHAVHAFITKEKDYVYQDFGRIEDIEKTFKDVFDYDGKYSRYR 352
Db 376 mgwmndmlkymetppyerhrhngvtfsllays-----enfilpshdevvhgkks--- 427
Qy 353 GRTHGAPVGDLPFRKFVFIQNHQVGNRNG---ERLSILTDTKTYVLMATLYILSPYI 409
Db 428 -----linkmpgsyeekfaqlrllygymma-----hpgk 456
Qy 410 PLIFMGEEYETNPFDFSDPVLKGVREGRLKENNOMIDPQSEAEFLSKLSW--- 466
Db 457 klifmgnef-----aqfdewkfedeldwvlf 482
Qy 467 --KIDEVLDYKOLINIRKRY 486
Db 483 dfeihrkmdymkeliacykry 504

RESULT 13
ID AAR96109 standard; Protein; 652 AA.
XX AAR96109;
XX 20-AUG-1996 (first entry)
XX Starch-branching-enzyme.
KW Starch-branching-enzyme; thermostable; Q-enzyme; Escherichia coli;
KW starch; amylopectin; glucan; cyclodextrin; food; food-additive;
KW adhesive; biodegradable plastic.

XX Bacillus stearothermophilus strain TRBE14 (FERM P-13916).
XX Key Location/Qualifiers
FH Peptide 238..249
FT /note= "Conserved amylase sequence used to
FT construct primer AAT27126"
FT Peptide 309..317
FT /note= "Conserved amylase sequence"
FT Peptide 414..423
FT /note= "Conserved amylase sequence used to
FT construct primer AAT27127"
XX EP710674-A2.
XX 08-MAY-1996.
XX 13-SEP-1995; 95EP-0250222.
XX 31-JUL-1995; 95JP-0195647.
XX 13-SEP-1994; 94JP-0218554.
XX (EZAK) EZAKI GLICO CO LTD.
XX Fujii K, Imanaka T, Nakamura H, Okada S, Takaha T;
PI Takata H, Terada Y, Yanase M;
XX WPI: 1996-223407/23.
XX N-PSDB; AAT27128.
XX New glucan with cyclic, branched core and outer branched region -
PT useful in foods, drinks, adhesives, infusion compns. etc., readily
PT soluble in water forming solns. of low viscosity
XX Example A-2; Page 28-31; 50pp; English.

CC This sequence represents a thermostable starch-branching-enzyme
CC (Q-enzyme, EC-2.4.1.18) from Bacillus stearothermophilus TRBE14
CC (FERM P-13916). Conserved regions present in amylases are used to
CC construct primers AAT27126-7, which are used in PCR to amplify a
CC fragment of the gene, which is then used as a probe to isolate a
CC full-length clone in Escherichia coli. The recombinant enzyme may
CC be reacted with starch or amylopectin to produce a glucan with a
CC degree of polymerisation of at least 50, with an inner branched
CC cyclic region and an outer branched region. The product is freely
CC soluble in water, forming a low-viscosity solution, is more
CC resistance to degradation than native starch, does not react with
CC proteins and amino acids, and is degraded to glucose in vivo, so is
CC easily digestible. The glucan may be used in foods, beverages,
CC food-additives, adhesives, in cyclodextrin production and as a
CC starch substitute in biodegradable plastic.
XX Sequence 652 AA;
SQ
Query Match 12.2%; Score 360; DB 17; Length 652;
Best Local Similarity 25.3%; Pred. No. 1.7e-24;
Matches 142; Conservative 72; Mismatches 186; Indels 162; Gaps 25;
Qy 4 FGNIEKNKG----IFKLWAPYVNSVKL-----KLSKKLIPMEKNDGFFFEVE 47
Db 26 fgahvineggkvgrfcwqapharevrlvgfsfndwdgtfdrlek-----vndegvwtiv 79
Qy 48 IDDEENLTYSY-IIEDKREI---PDPAARYQPLGVHDKSOLIRTD---YQILDLG---- 96
Db 80 vpenleghlykyeivtpdgqvlfkadpyafyselrptas---laydlkgvqndqswkrk 137
Qy 97 ----KVKIEDLIIVELHVGTF-SQEGNFKGVIEKLD----YKLDLGTITGELMPVAFPG 147
Db 138 krrkriydgqmvlyellhfgskkkdgrfityremadelisvldhgthiellplvehpl 197
Qy 148 NRWDGYDGVFLYAVONTYGGPWEKALVNEAHKRGIAVILDVYVNHII--GPEGNYLLGLG 205
Db 198 drswgyggtgyyavtsrygtphdfmyfvdrcdqagigvmdwvpgfhfckdahglymfdga 257
Qy 206 PYF-----SDRYKTPWGLTFNFDGDDQVRKFFLENVEYWFVKFKIDGLRLDAV----- 255
Db 258 ptveyanekdrenyvwg-tanf-dlgkpevrsflisnalfwleyhldgfrvdavnmly 315
Qy 256 ----HAIFDNS-PKHILQEIAEK--AHOLGKFVIAESDLNDPKIVKD-----DCGYKID 302
Db 316 wpndrilyenpyaveflrkineavfaydpnalmiaedstdwkvptaptgyeglgfnykwn 375
Qy 303 AQWVDDF-----HHAVHAFITKEKDYVYQDFGRIEDIEKTFKDVFDYDGKYSRYR 352
Db 376 mgwmndmlkymetppyerhrhngvtfsllays-----enfilpshdevvhgkks--- 427
Qy 353 GRTHGAPVGDLPFRKFVFIQNHQVGNRNG---ERLSILTDTKTYVLMATLYILSPYI 409
Db 428 -----linkmpgsyeekfaqlrllygymma-----hpgk 456
Qy 410 PLIFMGEEYETNPFDFSDPVLKGVREGRLKENNOMIDPQSEAEFLSKLSW--- 466
Db 457 klifmgnef-----aqfdewkfedeldwvlf 482
Qy 467 --KIDEVLDYKOLINIRKRY 486
Db 483 dfeihrkmdymkeliacykry 504
RESULT 14
ID AAR23787 standard; Protein; 658 AA.
XX AAR23787;
XX 04-NOV-1992 (first entry)
XX Heat-resistant pullanase.

XX Enzyme; pTPI.
XX Synthetic.
OS JP04099489-A.
PN 31-MAR-1992.
XX 27-MAR-1990; 90JP-0077230.
XX 27-MAR-1990; 90JP-0077230.
XX (EZAK) EZAKI GLICO CO.
PA (IATR) IATRON LAB INC.
XX WPI; 1992-157263/19.
DR P-PSDB; AAR23787.
XX Heat-resistant pullulanase gene - present recombinant plasmid are new
PT and produced in large ams. using a proper host
PS Claim 3; Page 2-3; 8pp; Japanese.
XX The sequence encoding this protein may be contained in a recombinant
CC plasmid. The heat-resistant pullulanase can be produced in large ams.
CC using a suitable host.
CC Plasmid pTPI contg. the pullulanase gene (Japan Patent Apppln. No.
CC 174037/88) was cleaved by various restriction enzymes and inserted
CC into plasmid pV118/119. Single-stranded DNA was prep'd. using helper
CC phage M13K07 and the base sequence determined by the dideoxy method.
XX Sequence 658 AA;

Query Match 12.1%; Score 356; DB 13; Length 658;
Best Local Similarity 22.9%; Pred. No. 4.1e-24;
Matches 141; Conservative 99; Mismatches 200; Indels 176; Gaps 26;

QY 2 FSPGGNI-----EKNKIGIFKLWAPVNSVKLKL-----SKKLIPMEKNDGFEVEIDDI 51
DB 44 ffygdgldvaytrekifkfwaptatevkvklidptktaeyalalerlekvgvttvtfgd 103

QY 52 EENLAYSII---EDKREIPDPASRYQPLGVHDKSLIRTDY-QIIDLKGVKIE----- 101
DB 104 megrytyvvcnavvreaavp-----yavavsvngeygiidlakrttekpipp 154

QY 102 -----DLIYELHVGTF-----SQEGNFKGVIEK-----LDYKDLGITGI 137
DB 155 itspadaviyemhirdftihpesgvthkgylgltetpetkgnptptglsylkelgvtvh 214

QY 138 ELPVQAQFPG-----NRDWCYDGVFLYAVQNTY---GGPW---ELAKLVNEAKHR 191
DB 215 ellpfndfagvdrnplqynvgynphynapegsyatpndpnyarirelkrailtqlsn 274

QY 182 GTAVILDVYNNHIGPEGNYL-----LGLGYFS-DRYKTPWGLTFNFDRCDCQ--V 230
DB 275 girvmdvynhv-----yirdqssfekivpvyfyfaygpnsgtvgndiaserkmv 329

QY 231 RFEILENVEYWFKFIDGLRLDAVIAIFDNPSPKHILQIEAKAHLGKFVIAESLNDP 290
DB 330 rkwidsvrfweeyhvngfrfdlm-gildvetmkavreml-----dtdlp 374

QY 291 KIVKDCGVKIDAQWDDFHVAHVAITREKDYIYDFGRIEDIEKTFKDFVYDCKYSR 350
DB 375 slifvgefwdlptp-----isseqkatmqnaeklprl-gyfnidyfrdsykgst 421

QY 351 YRGTRHGAPVGD-----LPPRFVVFIONHD-----OVGN 380
DB 422 fqlpdrfgalgsddrekvkaigsigknglflhtptqinyveshdhntfwdkmeian 481

QY 381 RGNGERLSILTDTKTYLMAATLYLSPIPLIFMGEYYETNPFFFDSPDLVLIKGYR 440
DB 155 GFLYAVQNTYGGPWELAKLVNEAKHRGTAIVLDVYNNH---GPEGNYLLGLGPFY---- 208

DB 482 gheseq----trkrqklatavllsqgipflhsgqefyrtk-----qgi- 522
QY 441 EGRLENNQNMIDPOSEEAFLKSKLSWKIDVELD---YVKOLINIRKRYNNCKRVEVR 497
DB 523 -----dmsynapds-----inqidwkrksehdvrvyvgllirkrkshraffttaei 571
QY 498 EGNCTILIMEKIGI 513
DB 572 nnhlffleptasia 587

RESULT 15
AAR11271
ID AAR11271 standard; Protein; 639 AA.
XX AAR11271;
XX 29-MAY-1991 (first entry)
XX B.stearothermophilus branching enzyme.
XX Branching enzyme; amylaceous substances; transferase;
XX thermostability.
XX Bacillus stearothermophilus.
XX EP418945-A.
XX 27-MAR-1991.
XX 22-AUG-1990; 90EP-0202264.
XX 23-AUG-1989; 89NL-0002128.
XX (COOP-) COOP VERKOOP AVEBE.
XX Kiel JAKW, Boels JM, Beldman G, Venema G;
XX WPI; 1991-088587/13.
XX N-PSDB; AAQ11072.
XX Branching enzyme obt'd. from Bacillus stearothermophilus - has good
XX thermostability, for use in modifying starch-like substances
XX Disclosure; fig 4; 13pp; English.
XX This branching enzyme is derived from B.stearothermophilus and is
XX encoded by the gigB gene. It is stable at high temps. and can be
XX produced in large ams. using recombinant techniques. It is a
XX transferase allowing formation of alpha(1-6) branches in starch-like
XX substances. It is useful for extending the shelf life of such
XX materials and can be used in the prepn. of human and animal foods.
XX Sequence 639 AA;

Query Match 11.6%; Score 342.5; DB 12; Length 639;
Best Local Similarity 24.9%; Pred. No. 7.2e-23;
Matches 135; Conservative 89; Mismatches 197; Indels 121; Gaps 26;

QY 4 FGNIEKNKI-----FKLWAPVNSVKLKL-----KKLIPMEKNDGFEVEIDIEE 53
DB 26 fgahvikkngmvgtrfcvwapsharevrlvgsfnewngtnfmlmkvsnqgvwmifipenle 85

QY 54 NLTYSVII---EKREI---PDPASRYQPLGVHDKSQLIR-TDYQILD-----LGKVKI 100
DB 86 ghlykyeitndgndvnlksdpafyafyselrphtasivynikgyqndqtwrrkkqrkriyd 145

QY 101 EDLIIVELHVGTF--SQEGNF---KGVIEKLDYKDLGITGIELMPVAQFPCNRDMGYD 154
DB 146 qplfiyghfsgwkkkedgsfityqemaeeilpyviehgfthiellplvehpdrswgyq 205

QY 155 GFLYAVQNTYGGPWELAKLVNEAKHRGTAIVLDVYNNH---GPEGNYLLGLGPFY---- 208

```

Db      206 gigyysatsrygtbhdhmyfidrchqagigvllgwpgphfckdshgylmfdgapayeyan 265
Qy      209 -SDRYKTPWGLTFNFDORGCQVRKFILFENVEYWFKFKIDGLRLDAV-----HAI 258
Db      266 mqdrenywg-tanf-dlgpevrsflisnalfwmevfhdgfrvdavanmlywpnsdvl 323
Qy      259 PDNS-PKHILQEIAEK--AHOLGKFVIAESDLNDPKIVKD--DCG----YKIDAQWVDDF 309
Db      324 ykntyaveflqklnetvfaydpnlmliaedstdwprvtaptgydgglgfnykwnmgwmndi 383
Qy      310 HHAVHAFITTEKDYIYODFGRIEDIEKTFKDVVYDGGKYSRYRCRTHGAPVGDLPPrKFV 369
Db      384 --ltymetppehrkyvhn-----kvtfsllyaysef----- 413
Qy      370 VFIONHQVGNRNGERLSIL---TDKTTYLMAATLYILS-PYIPLIFMGEEYETNPf 424
Db      414 ilpfshdevh-gkksllskmpgtyeekfaqlrillygylthpghkklfmggefgq---- 468
Qy      425 FFFSDFSDPVLKGVREGRLKENNOMIDPOSEEAFKLSKLSWKIDDEVLDYIKOLINIRK 484
Db      469 --fdewkd-----leqldwml-----fdmhrnmmyvkvellkcyk 503
Qy      485 RY 486
Db      504 ry 505

```

Search completed: July 15, 2002, 12:17:34
 Job time: 71 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 12:18:04 ; Search time 25.09 Seconds
(without alignments)
541.277 Million cell updates/sec

Title: US-09-298-924-8
Perfect score: 2954
Sequence: 1 MFSFGNIEKNKGIKLMAP.....KLKDELIKVNRGVGVYQLE 556

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2954	100.0	556	US-08-505-377-1	Sequence 1, Appli
2	2954	100.0	556	US-08-798-269-1	Sequence 1, Appli
3	2954	100.0	556	US-09-055-210-1	Sequence 1, Appli
4	1748.5	59.2	559	US-09-242-690A-15	Sequence 15, Appli
5	969	32.8	589	US-08-399-646-2	Sequence 2, Appli
6	969	32.8	589	US-08-607-321-2	Sequence 2, Appli
7	969	32.8	589	US-08-961-240-2	Sequence 2, Appli
8	969	32.8	589	US-08-603-501-2	Sequence 2, Appli
9	969	32.8	596	US-08-399-646-12	Sequence 12, Appli
10	969	32.8	596	US-08-607-321-12	Sequence 12, Appli
11	969	32.8	596	US-08-961-240-12	Sequence 12, Appli
12	969	32.8	596	US-08-605-501-12	Sequence 12, Appli
13	959	32.5	597	US-08-399-646-4	Sequence 4, Appli
14	959	32.5	597	US-08-607-321-4	Sequence 4, Appli
15	959	32.5	597	US-08-961-240-4	Sequence 4, Appli
16	959	32.5	597	US-08-605-501-4	Sequence 4, Appli
17	959	32.5	598	US-08-399-646-14	Sequence 14, Appli
18	959	32.5	598	US-08-607-321-14	Sequence 14, Appli
19	959	32.5	598	US-08-961-240-14	Sequence 14, Appli
20	959	32.5	598	US-08-603-501-14	Sequence 14, Appli
21	360	12.2	652	US-08-528-026C-4	Sequence 4, Appli
22	319.5	10.8	718	US-09-346-237-10	Sequence 10, Appli
23	310.5	10.5	606	US-09-187-124-2	Sequence 2, Appli
24	300.5	10.2	740	US-08-410-784A-4	Sequence 4, Appli
25	298	10.1	713	US-09-346-237-9	Sequence 9, Appli
26	296.5	10.0	818	US-08-410-784A-2	Sequence 2, Appli
27	296.5	10.0	818	US-09-346-237-11	Sequence 11, Appli

28	292	9.9	829	4	US-09-514-599-6	Sequence 6, Appli
29	292	9.9	862	4	US-09-346-237-1	Sequence 1, Appli
30	284	9.6	893	4	US-09-514-302-4	Sequence 4, Appli
31	284	9.6	1938	4	US-09-514-302-2	Sequence 2, Appli
32	276	9.3	921	4	US-09-514-599-2	Sequence 2, Appli
33	268	9.1	928	4	US-09-346-237-2	Sequence 2, Appli
34	268	9.1	915	4	US-09-514-599-4	Sequence 4, Appli
35	264.5	9.0	820	4	US-09-313-677-21	Sequence 21, Appli
36	264.5	9.0	926	4	US-09-313-677-2	Sequence 2, Appli
37	264.5	9.0	933	4	US-09-313-677-19	Sequence 19, Appli
38	264.5	9.0	967	4	US-09-313-677-17	Sequence 17, Appli
39	250	8.5	928	1	US-08-474-140-11	Sequence 11, Appli
40	250	8.5	928	1	US-08-477-630-11	Sequence 11, Appli
41	250	8.5	928	1	US-08-472-293-11	Sequence 11, Appli
42	250	8.5	928	1	US-08-474-545-11	Sequence 11, Appli
43	250	8.5	928	2	US-08-478-341-11	Sequence 11, Appli
44	250	8.5	928	3	US-08-996-733-11	Sequence 11, Appli
45	249	8.4	588	2	US-08-339-715A-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-08-505-377-1
; Sequence 1, Application US/08505377
; Patent No. 5856146
; GENERAL INFORMATION:
; APPLICANT: MITSUZUMI, Hitoshi
; APPLICANT: KUBOTA, Michio
; APPLICANT: SUGIMOTO, Toshiyuki
; TITLE OF INVENTION: RECOMBINANT THERMOSTABLE ENZYME WHICH
; TITLE OF INVENTION: RELEASES TREHALOSE FROM NON-REDUCING SACCHARIDE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Browdy and Neimark
; STREET: 419 Seventh Street N.W. Ste. 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/505,377
; FILING DATE: 21-JUL-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 190180/1994
; FILING DATE: 21-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 109128/1995
; FILING DATE: 21-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP NOT YET RECEIVED
; FILING DATE: 04-JUL-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Browdy, Roger L
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: MITSUZUMI-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-5197
; TELEFAX: (202) 737-3528
; TELEX: 249688
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 556 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein

Claims copy pep.

US-08-505-377-1

Query Match 100.0%; Score 2954; DB 2; Length 556;
Best Local Similarity 100.0%; Pred. No. 1.1e-280;
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MFSGGNIENKNGIFKLWAPVNSVKLKLIPMEKNDGEGFEVEIDDEENLTYSYI 60
Db 1 MFSGGNIENKNGIFKLWAPVNSVKLKLIPMEKNDGEGFEVEIDDEENLTYSYI 60

Qy 61 IEDKREIPDPASRYOPLGVHDKSQILRTDYQILDGKVKIEDLIYELHVGTFSSQENFK 120
Db 61 IEDKREIPDPASRYOPLGVHDKSQILRTDYQILDGKVKIEDLIYELHVGTFSSQENFK 120

Qy 121 GVIEKLDYLDGLGTITELMPVAQPGNRDWDYGVFLYAVQNTYGGPWELAKLVNEAHK 180
Db 121 GVIEKLDYLDGLGTITELMPVAQPGNRDWDYGVFLYAVQNTYGGPWELAKLVNEAHK 180

Qy 181 RGIAVILDVYVNHIGPEGNYLLGLGPFSDRYKTPWGLTFNFDGCDQVRKFLENVEY 240
Db 181 RGIAVILDVYVNHIGPEGNYLLGLGPFSDRYKTPWGLTFNFDGCDQVRKFLENVEY 240

Qy 241 WFTFKIDGLRLDAVHAIFDNPSPKHILQETAIAKAHQLGKGFVIAESDNDPKIVKDDCGYK 300
Db 241 WFTFKIDGLRLDAVHAIFDNPSPKHILQETAIAKAHQLGKGFVIAESDNDPKIVKDDCGYK 300

Qy 301 IDAQWDDFHAVHAFITKEKDYQDFGRIEDIEKTFKDVYVYDGYKSYRGRTHGAPV 360
Db 301 IDAQWDDFHAVHAFITKEKDYQDFGRIEDIEKTFKDVYVYDGYKSYRGRTHGAPV 360

Qy 361 IDLPKPRKRVFVQIKNHDOVGNRGNGERLSILTDKTYLMAATLYLSPIPLFMGEYEE 420
Db 361 IDLPKPRKRVFVQIKNHDOVGNRGNGERLSILTDKTYLMAATLYLSPIPLFMGEYEE 420

Qy 421 TNPPFFSDFSDPVLKGVREGRLKENNOMIDPOSEAFKLSKSWKIDDEVLDYKOLI 480
Db 421 TNPPFFSDFSDPVLKGVREGRLKENNOMIDPOSEAFKLSKSWKIDDEVLDYKOLI 480

Qy 481 NIKRYNNKRVKVRREGNCITLIMEKIGIIASFDDIVINSKITGNLLIGIGPPKLLK 540
Db 481 NIKRYNNKRVKVRREGNCITLIMEKIGIIASFDDIVINSKITGNLLIGIGPPKLLK 540

RESULT 2
US-08-798-269-1
; Sequence 1, Application us/08798269
; Patent No. 6027918
; GENERAL INFORMATION:
; APPLICANT: MITSUZUMI, Hitoshi
; APPLICANT: KUBOTA, Michio
; APPLICANT: SUGIMOTO, Toshiyuki
; TITLE OF INVENTION: RECOMBINANT THERMOSTABLE ENZYME WHICH
; TITLE OF INVENTION: RELEASES TREHALOSE FROM NON-REDUCING SACCHARIDE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Browdy and Neimark
; STREET: 419 Seventh Street N.W. Ste. 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/798,269

Handwritten: **NAME DNA**

FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/505,377
FILING DATE: 21-JUL-1995
APPLICATION NUMBER: JP 190180/1994
FILING DATE: 21-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 109128/1995
FILING DATE: 11-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP NOT YET RECEIVED
FILING DATE: 04-JUL-1995
ATTORNEY/AGENT INFORMATION:
NAME: Browdy, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: MITSUZUMI-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-5197
TELEFAX: (202) 737-3528
TELEX: 249688
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 556 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-798-269-1

Query Match 100.0%; Score 2954; DB 3; Length 556;
Best Local Similarity 100.0%; Pred. No. 1.1e-280;
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MFSGGNIENKNGIFKLWAPVNSVKLKLIPMEKNDGEGFEVEIDDEENLTYSYI 60
Db 1 MFSGGNIENKNGIFKLWAPVNSVKLKLIPMEKNDGEGFEVEIDDEENLTYSYI 60

Qy 61 IEDKREIPDPASRYOPLGVHDKSQILRTDYQILDGKVKIEDLIYELHVGTFSSQENFK 120
Db 61 IEDKREIPDPASRYOPLGVHDKSQILRTDYQILDGKVKIEDLIYELHVGTFSSQENFK 120

Qy 121 GVIEKLDYLDGLGTITELMPVAQPGNRDWDYGVFLYAVQNTYGGPWELAKLVNEAHK 180
Db 121 GVIEKLDYLDGLGTITELMPVAQPGNRDWDYGVFLYAVQNTYGGPWELAKLVNEAHK 180

Qy 181 RGIAVILDVYVNHIGPEGNYLLGLGPFSDRYKTPWGLTFNFDGCDQVRKFLENVEY 240
Db 181 RGIAVILDVYVNHIGPEGNYLLGLGPFSDRYKTPWGLTFNFDGCDQVRKFLENVEY 240

Qy 241 WFTFKIDGLRLDAVHAIFDNPSPKHILQETAIAKAHQLGKGFVIAESDNDPKIVKDDCGYK 300
Db 241 WFTFKIDGLRLDAVHAIFDNPSPKHILQETAIAKAHQLGKGFVIAESDNDPKIVKDDCGYK 300

Qy 301 IDAQWDDFHAVHAFITKEKDYQDFGRIEDIEKTFKDVYVYDGYKSYRGRTHGAPV 360
Db 301 IDAQWDDFHAVHAFITKEKDYQDFGRIEDIEKTFKDVYVYDGYKSYRGRTHGAPV 360

Qy 361 IDLPKPRKRVFVQIKNHDOVGNRGNGERLSILTDKTYLMAATLYLSPIPLFMGEYEE 420
Db 361 IDLPKPRKRVFVQIKNHDOVGNRGNGERLSILTDKTYLMAATLYLSPIPLFMGEYEE 420

Qy 421 TNPPFFSDFSDPVLKGVREGRLKENNOMIDPOSEAFKLSKSWKIDDEVLDYKOLI 480
Db 421 TNPPFFSDFSDPVLKGVREGRLKENNOMIDPOSEAFKLSKSWKIDDEVLDYKOLI 480

Qy 481 NIKRYNNKRVKVRREGNCITLIMEKIGIIASFDDIVINSKITGNLLIGIGPPKLLK 540
Db 481 NIKRYNNKRVKVRREGNCITLIMEKIGIIASFDDIVINSKITGNLLIGIGPPKLLK 540

Qy 541 DELIKVNRGVGYQLE 556
Db 541 DELIKVNRGVGYQLE 556

RESULT 3
US-09-055-210-1
; Sequence 1, Application US/09055210
; Patent No. 6346394
; GENERAL INFORMATION:
; APPLICANT: MITSUZUMI, Hitoshi
; APPLICANT: KUBOTA, Michio
; APPLICANT: SUGIMOTO, Toshiyuki
; TITLE OF INVENTION: RECOMBINANT THERMOSTABLE ENZYME WHICH
; RELEASES TREHALOSE FROM NON-REDUCING SACCHARIDE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Browdy and Neimark
; STREET: 419 Seventh Street N.W. Ste. 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; -COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/055,210
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/505,377
; FILING DATE: 21-JUL-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 109128/1995
; FILING DATE: 11-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP NOT YET RECEIVED
; FILING DATE: 04-JUL-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Browdy, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: MITSUZUMI-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-5197
; TELEFAX: (202) 737-3528
; TELEX: 249688
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 556 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-055-210-1

Query Match 100.0%; Score 2954; DB 4; Length 556;
Best Local Similarity 100.0%; Pred. No. 1.1e-280;
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MFSFGNIEKNKGIFKLPAPYNSVKLSKLLIPMEKNDGEGFEVEIDIEENLTYSYI 60
DB 1 MFSFGNIEKNKGIFKLPAPYNSVKLSKLLIPMEKNDGEGFEVEIDIEENLTYSYI 60
QY 61 IEDKREIPDPASRYQPLGVHDKSQLRTDYQLDLGKVKIEDLIYELHVGTFSQGNEK 120
DB 61 IEDKREIPDPASRYQPLGVHDKSQLRTDYQLDLGKVKIEDLIYELHVGTFSQGNEK 120
QY 121 GVIEKLDYLDLIGITGIELMPVAQPGNPDWGYDGVFLYAVQNTYGGPWELAKLVNEAHK 180
DB 121 GVIEKLDYLDLIGITGIELMPVAQPGNPDWGYDGVFLYAVQNTYGGPWELAKLVNEAHK 180
QY 181 RGIADVLDVYVNHIGPEGNYLLGLSPYFSDRYKTPMGLTFNFDRCQDQVRKFIENVEY 240

DB 181 RGIADVLDVYVNHIGPEGNYLLGLSPYFSDRYKTPMGLTFNFDRCQDQVRKFIENVEY 240
QY 241 WFKTEFKIDGLRLDAVHAIFDNPSPKHLOEIAEKAHOLGKFVIAESDLNDPKIVKDDCGYK 300
DB 241 WFKTEFKIDGLRLDAVHAIFDNPSPKHLOEIAEKAHOLGKFVIAESDLNDPKIVKDDCGYK 300
QY 301 IDAQWVDDFHHAHAFITKEKDYQDFGRIEDIEKTFKDVYDYGKYSRYGRTHGAPV 360
DB 301 IDAQWVDDFHHAHAFITKEKDYQDFGRIEDIEKTFKDVYDYGKYSRYGRTHGAPV 360
QY 361 GDLPPRKFVVFVIONHDOVNGRNGERLSILTDKTTVYMAATLYLSPYPLIFPMGEYIE 420
DB 361 GDLPPRKFVVFVIONHDOVNGRNGERLSILTDKTTVYMAATLYLSPYPLIFPMGEYIE 420
QY 421 TNPFFFFSDFSDPVLKIGVREGRLKENNOMIDPQSEAFKLSKLNKIDDEVLDYKQLI 480
DB 421 TNPFFFFSDFSDPVLKIGVREGRLKENNOMIDPQSEAFKLSKLNKIDDEVLDYKQLI 480
QY 481 NIKRYNNCKRKVEVRREGNCITLIMEKIGIIASFDDIVINSKITGNLLIGIGFPKKLKK 540
DB 481 NIKRYNNCKRKVEVRREGNCITLIMEKIGIIASFDDIVINSKITGNLLIGIGFPKKLKK 540
QY 541 DELIKVNRGVGYQLE 556
DB 541 DELIKVNRGVGYQLE 556

RESULT 4
US-09-242-690A-15
; Sequence 15, Application US/09242690A
; Patent No. 6284534
; GENERAL INFORMATION:
; APPLICANT: KONDO, KEIJI
; APPLICANT: MIURA, YUTAKA
; TITLE OF INVENTION: YEAST VECTOR AND METHOD OF PRODUCING PROTEINS USING THE
; FILE REFERENCE: 049441/0118
; CURRENT APPLICATION NUMBER: US/09/242,690A
; PRIOR FILING DATE: 1999-02-23
; PRIOR APPLICATION NUMBER: PCT/JP97/02924
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: JP 8/241062
; PRIOR FILING DATE: 1996-08-23
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 559
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic protein
US-09-242-690A-15

Query Match 59.2%; Score 1748.5; DB 4; Length 559;
Best Local Similarity 58.9%; Pred. No. 1.3e-162;
Matches 329; Conservative 87; Mismatches 136; Indels 7; Gaps 5;
QY 3 SFGGNIENKKGIFKLPAPYNSVKLK-LSKKLLIPMEKNDGEGFEVEIDIEENLTYSYI 61
DB 2 TAYKIDGNEVITPLWAPYQSVKLVKGLYEMERDEKGYTFTINNVKVDRYKYL 61
QY 62 EDKREIPDPASRYQPLGVHDKSQLRTDYQLDLGKVKIEDLIYELHVGTFSQGNEK 121
DB 62 EDKREIPDPASRYQPLGVHDKSQLRTDYQLDLGKVKIEDLIYELHVGTFSQGNEK 121
QY 122 GVIEKLDYLDLIGITGIELMPVAQPGNPDWGYDGVFLYAVQNTYGGPWELAKLVNEAHK 181
DB 122 GVIEKLDYLDLIGITGIELMPVAQPGNPDWGYDGVFLYAVQNTYGGPWELAKLVNEAHK 181
QY 182 RGIADVLDVYVNHIGPEGNYLLGLSPYFSDRYKTPMGLTFNFDRCQDQVRKFIENVEY 241
DB 182 RGIADVLDVYVNHIGPEGNYLLGLSPYFSDRYKTPMGLTFNFDRCQDQVRKFIENVEY 241

```
QY 242 FRTFKIDGLRDADVAHAIIDNSPKHILQIEIAEKAHQKGFVIAESDLNDPKIV--KDCG 299
| : ||| ||| ||| ||| ||| : | : ||| ||| ||| : | : |||
Db 242 IKEYNVDFGLRDADVAHAIIDTSPKHILEIADVVHKNRIVIAESDLNDPRVNPKEKCG 301
| : ||| ||| ||| ||| ||| : | : ||| ||| ||| : | : |||
QY 300 KIDAQWDDFHHAHVAFITKEKDYQDFGRIEDIEKTFKDFVYDGYKSYRGRTHGAP 359
||||||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db 302 NIDAQWDDFHHSIHAYLTGERQYDFDGNLDDIVKSYKDFVYDGYKSYNFRKTHGEP 361
||||||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
QY 360 VGLDPPRKVFVFIQNHQVGNRNGERLSITLTDKTTVYMAATLYILSPYIPLFMGEYY 419
||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db 362 VGLDGCNFFVYIIONHDOVGNRNGERLIILVDRESYKIAAALYLLSPYIMPFMGEYG 421
| ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
QY 420 ETNPFYFSDFSDPVLKIGVREGRLKENNQMDPQSEAEFLKSLKSWKIDEEVLDYKQL 479
| ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db 422 EENPFYFSDFSKSLQGVREGRLKENQMDPQSEAEFLKSLKSWKIDEEIFSKYL 481
| ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
QY 480 INIRKRYN-NCKRVKVEYRNGNCITLMEKIGIITASFDDIVINSKITGNLLIGI--GFPK 536
| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db 482 IKMKELSIACDRRVNVNGENMLIKGREYFSLYVFSKSSIEVKYSGTLLLSNNSFPQ 541
| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
QY 537 KLKDELKIVNRGVGVYQL 555
| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db 542 HIEEGK-YEFDKGFALYKL 559
| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||

RESULT 5
US-08-399-646-2
; Sequence 2, Application US/08399646
; Patent No. 5556781
; GENERAL INFORMATION:
; APPLICANT: KUBOTA, Michio
; APPLICANT: TSUSAKI, Keiji
; APPLICANT: HATTORI, Kazuko
; APPLICANT: SUGIMOTO, Toshiyuki
; TITLE OF INVENTION: DNA ENCODING ENZYME, RECOMBINANT DNA AND
; TITLE OF INVENTION: ENZYME, TRANSFORMANT, AND THEIR PREPARATIONS AND USE
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/399,646
; FILING DATE: 07-MAR-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 59834
; FILING DATE: 07-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 59840
; FILING DATE: 07-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: KUBOTA-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 589 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
```

```
; MOLECULE TYPE: protein
US-08-399-646-2
Query Match 32.8%; Score 969; DB 1; Length 589;
Best Local Similarity 39.5%; Pred. No. 3.2e-86;
Matches 215; Conservative 74; Mismatches 195; Indels 60; Gaps 12;

QY 13 GIKFLWAPYVNSVKLSKSLIPMEKN-----DEGFEVEIDIEENLYSYIE-DKR 65
| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db 8 GRDIWAPEAGTVTLLAGGERYEMGRPRGNGPADEGHWTAADAPTGAADVGYLLDGEI 67
| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
QY 66 EIPDPASRYOPLGVHDKSQLIRTDYQILDGLKVKIEDL-----IYELHVGTFQS 115
| ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db 68 PLDPRTRRRQPEGVH---ALSRT----FDFGAHRWQDAGWQRELOGSVIVELHIGTFP 120
| ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
QY 116 EGNFKGVIEKLDYKDLGITGIELMPVAQFPNGRNDWGYDGVFLYAVONTYGGWELAKV 175
| ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db 121 EGTIDAAAGKLDYLAGLIDFIELLPYNAFNTHNMGYDGVQWFAVHEGYGGAAYQRFV 180
| ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
QY 176 NEAHKRGIAVILDVYVNHIGPEGNLYLLGLPYFSDRYKTPMGLTFNDDRGCDQVRKFI 235
| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db 181 DAAHAAGLVIQDVVYNHLGSPGNLYLPYGPYKLGEGNTWGDVNLDDGPGSDHVRQYIL 240
| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
QY 236 ENVEYMFKTEKIDGLRDADVAHAIIDNSPKHILQIEIAEKAHQKGFVIAESDLND 289
| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db 241 DNVAMLRDYRVDGLRDADVAHAIIDNSPKHILQIEIAEKAHQKGFVIAESDLND 300
| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
QY 290 PKIV--KDCGKYKIDAQWDDFHHAHVAFITKEKDYQDFGRIEDIEKTFKDFVYD 347
| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db 301 PRLLYPRDNGYGLAGQWSDDFHHAHVNVSGTGTGYSDFLSALAKVLRDGFHDS 360
| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
QY 348 YSRYRGRTHGAPV--GDLPPRKVFVFIQNHQVGNRNGERLSITLTDKTTVYMAATLY 405
| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db 361 YSFRGRCHRPINFSAHVAAALVVCVSNHDOIGNRATGDRLSQSLPYGSLAAVTLT 420
| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
QY 406 SPYIPLFMGEYYETNPFYFSDFSDPVLKIGVREGRLKENNQ-----IDPQSEEA 458
| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db 421 GPFTPLFMGEYGAATTQWQFTSHPELKGATAGRIREFERMGMWDPAVVPDPQPET 480
| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
QY 459 FLASKLSWK-----IDDEVLDYKQLINIRKRYNCKRVKVEYRNGNCITLMEKIGI 513
| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db 481 FTRSKLDWAESASAGDHARLLLELYRSITLR-----RSTPELARLGPADTAV----- 526
| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
QY 514 SFDD 517
| |||
Db 527 EFDD 530

RESULT 6
US-08-607-321-2
; Sequence 2, Application US/08607321
; Patent No. 5716813
; GENERAL INFORMATION:
; APPLICANT: KUBOTA, Michio
; APPLICANT: TSUSAKI, Keiji
; APPLICANT: HATTORI, Kazuko
; APPLICANT: SUGIMOTO, Toshiyuki
; TITLE OF INVENTION: DNA ENCODING ENZYME, RECOMBINANT DNA AND
; TITLE OF INVENTION: ENZYME, TRANSFORMANT, AND THEIR PREPARATIONS AND USE
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
```

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/607,321
; FILING DATE: 26-FEB-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/399,646
; FILING DATE: 07-MAR-1995
; APPLICATION NUMBER: JP 59834
; FILING DATE: 07-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 59840
; FILING DATE: 07-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: KUBOTA-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 589 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-607-321-2

Query Match 32.8%; Score 969; DB 1; Length 589;
Best Local Similarity 39.5%; Pred. No. 3.2e-86;
Matches 215; Conservative 74; Mismatches 195; Indels 60; Gaps 12;

QY 13 GIFKLPVNSVKLSKLLIPMEKN-----DEGFEVEIDDEENTYSYVIE-DKR 65
DB 8 GRFDIWPAGVITLLAGERYEMRRPGNGPADEGWTAADAPTADVDYGYLLDGDDEI 67
QY 66 EIPDPASRQPLGVHDKSLIRTDYQILDGLKVKIEDL-----IYELHVGTFSSQ 115
DB 68 PLPDPTRRQPEGVH---ALSR-----FDPGAHRWQDAGWQRELOSGSVIYELHIGTFP 120
QY 116 EGNFGVIEKLDYLDLGTIGTIELMPVAFPGNDRNGYDGVFLYAVQNTYGGPWEAKLV 175
DB 121 EGTLDAAAGKLDYLAGLIGDIFELLPVNAFNGTHNNGYDGVQWFAVHEGYGPAAYQRFV 180
QY 176 NEAHKRGIAVILDVYVNHIGPEGNYLILGLGYPESDRYKTPWGLTFNFDGDCDOVRKFTL 235
DB 181 DAHAAGLVIOQVYVNHIGPSGNLYLPRYGPYLKHGEGNTWGSVNLDPGSDHVRQYIL 240
QY 236 ENVEYWFKTFKIDGLRDLAVHAFIFONSFKHILQETAEAKAHLGK-----FVIAESDLND 289
DB 241 DNVAMWLRDVRVYDGLRDLAVHAKDERAVHILEEFGALADALSSEGGRLTLIAESDLNN 300
QY 290 PKIV--KDCGKIDQWDDHVAHVHATITEKDYDYDFGRIEDIEKTFKDFVYDCK 347
DB 301 PRLLYPRDNGYGLAGQWDDFHVAHVHVVNSGETTGYSDFDSLALAKLVLRDGFPHDGS 360
QY 348 YSRYGRTHGAPV--GDLPRKRVFVIONHDOVNGRNGERLSILTDKTTYLMAATLYIL 405
DB 361 YSFRGRCHGRPINFSAVHAPALVCSQNHQIGNRATGDRLSQSLLPYGSLAAVLTIT 420
QY 406 SPYIPLIFMGEYETNPFFSDFSDPVLKIGVREGRLKENNQ-----IDPQSEEA 458
DB 421 GPPTPMLFNGEYGATTPWQFTSHPEPELGRKATASGRIFERMGWDPAVVPDQDPET 480
QY 459 FLKSKLSWK-----IDDEVLDYKQLINIRKYNCKRKEVRENCITLIMEKIGITA 513
DB 481 FTRSKLDWAEASAGDARLLELYRSILTLR-----RSTPELARLGFADTAV----- 526
QY 514 SFDD 517
DB 527 EFDD 530

; RESULT 7
; US-08-961-240-2
; Sequence 2, Application US/08961240
; Patent No. 5830715
; GENERAL INFORMATION:
; APPLICANT: KUBOTA, Michio
; APPLICANT: TSUSAKI, Keiji
; APPLICANT: HATTORI, Kazuko
; APPLICANT: SUGIMOTO, Toshiyuki
; TITLE OF INVENTION: DNA ENCODING ENZYME, RECOMBINANT DNA AND
; TITLE OF INVENTION: ENZYME, TRANSFORMANT, AND THEIR PREPARATIONS AND USE
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,240
; FILING DATE: 30-OCT-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/399,646
; FILING DATE: 07-MAR-1995
; APPLICATION NUMBER: JP 59834
; FILING DATE: 07-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 59840
; FILING DATE: 07-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: KUBOTA-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 589 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-961-240-2

Query Match 32.8%; Score 969; DB 2; Length 589;
Best Local Similarity 39.5%; Pred. No. 3.2e-86;
Matches 215; Conservative 74; Mismatches 195; Indels 60; Gaps 12;

QY 13 GIFKLPVNSVKLSKLLIPMEKN-----DEGFEVEIDDEENTYSYVIE-DKR 65
DB 8 GRFDIWPAGVITLLAGERYEMRRPGNGPADEGWTAADAPTADVDYGYLLDGDDEI 67
QY 66 EIPDPASRQPLGVHDKSLIRTDYQILDGLKVKIEDL-----IYELHVGTFSSQ 115
DB 68 PLPDPTRRQPEGVH---ALSR-----FDPGAHRWQDAGWQRELOSGSVIYELHIGTFP 120
QY 116 EGNFGVIEKLDYLDLGTIGTIELMPVAFPGNDRNGYDGVFLYAVQNTYGGPWEAKLV 175
DB 121 EGTLDAAAGKLDYLAGLIGDIFELLPVNAFNGTHNNGYDGVQWFAVHEGYGPAAYQRFV 180
QY 176 NEAHKRGIAVILDVYVNHIGPEGNYLILGLGYPESDRYKTPWGLTFNFDGDCDOVRKFTL 235
DB 181 DAHAAGLVIOQVYVNHIGPSGNLYLPRYGPYLKHGEGNTWGSVNLDPGSDHVRQYIL 240
```


OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/399,646
FILING DATE: 07-MAR-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 59834
FILING DATE: 07-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 59840
FILING DATE: 07-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: BROWDY, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: KUBOTA-5
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 596 amino acids
TYPE: amino acid
MOLECULE TYPE: protein
US-08-399-646-12

Query Match 32.8%; Score 969; DB 1; Length 596;
Best Local Similarity 39.5%; Pred. No. 3.2e-86;
Matches 215; Conservative 74; Mismatches 195; Indels 60; Gaps 12;

QY 13 GIFKLWAPYVNSVKLSKLLIPMEKN-----DEGFFEVEIDDIENLTYSYIIIE-DKR 65
DB 15 GRFDIWAPEAGTVTLGGERYEMGRPGNGPADEGWTAADAPTADVDYGYLLDGEI 74

QY 66 EIPDPASRYQPLGVHDKSLQIRTDYQILDGKVKIEDL-----IYELHVGTFSSQ 115
DB 75 PLPDPTRRQPEGVH---ALSRT---FDPGAHRWQDAGWQRELOGSVIYELHIGTFTP 127

QY 116 EGNFKGVIEKLDYKDLGTIGTIELMPVAQPPGNRNGYDGVFLYAVONTYGGPWELAKLV 175
DB 128 EGTDLAAAGKLDYLAGLIGDIFIELLPVNAFNGTHNGYDGVQWFAVHEGYGPPAAQYQRFV 187

QY 176 NEAHKRGIAVILDVYVNHGTPGNYLLGLPYFSDRYKTPWGLTFNFDRCGDOVKFKIL 235
DB 188 DAAHAAGLVQDVVYVNHGTPGNYLLPRYGPYKLGEGNTWGSVNLDPGSDHVRQYIL 247

QY 236 ENVEYWFKTFKIDGLRDVAHAFDINSKHIQLQEIKAHQGLK-----FVIAESDLND 289
DB 248 DNVAMWLRDVRVGLRDVAHAKDERAVHILEEFGALADALSSEGGRPPLTLIAESDLNN 307

QY 290 PKIV--KDCGKYIDAQWDDPHAVHAFITTEKQYIYDFGRIEDIEKTFKDVFYVDGK 347
DB 308 PRLLYPRDNGYGLAGQWSDDEPHAVHVNVSGETTGYSDFSGLAKALVLRDGFPHDGS 367

QY 348 YSRYRQTHGAPV--GDLPPRFVFIQNHQDQVNGRNGERLSILTDKTYTLMATLYIL 405
DB 368 YSFRGRCHRPINFSAVHPALVWCSQNHQDIGNRATGDRLSQSLPYSLAALAVLTLT 427

QY 406 SPYIPLIFMGEYEYENPFFSDFSDPVLKGVREGRKNNQM-----IDPQSEEA 458
DB 428 GFTPTMLFMEYEGATTPQOFFTSHPPELKGKATAGRIREFRGMWDPAVVPDPODPET 487

QY 459 FLKSKLSWK-----IDEEVDLYQKQLINIRKRYNNCKRYKVRREGNCITLIMEKIGIIA 513
DB 488 FTRSKLDWAEASAGDHARLLELYRSLLTLR-----RSTPELARLGFADTAV----- 533

QY 514 SFDD 517
DB 534 EFDD 537

RESULT 10
US-08-607-321-12
Sequence 12, Application US/08607321
Patent No. 5716813
GENERAL INFORMATION:
APPLICANT: KUBOTA, Michio
APPLICANT: TSUSAKI, Keiji
APPLICANT: HATTORI, Kazuko
APPLICANT: SUGIMOTO, Toshiyuki
TITLE OF INVENTION: DNA ENCODING ENZYME, RECOMBINANT DNA AND
TITLE OF INVENTION: ENZYME, TRANSFORMANT, AND THEIR PREPARATIONS AND USE
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/607,321
FILING DATE: 26-FEB-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/399,646
FILING DATE: 07-MAR-1995
APPLICATION NUMBER: JP 59834
FILING DATE: 07-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 59840
FILING DATE: 07-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: BROWDY, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: KUBOTA-5
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 596 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-607-321-12

Query Match 32.8%; Score 969; DB 1; Length 596;
Best Local Similarity 39.5%; Pred. No. 3.2e-86;
Matches 215; Conservative 74; Mismatches 195; Indels 60; Gaps 12;

QY 13 GIFKLWAPYVNSVKLSKLLIPMEKN-----DEGFFEVEIDDIENLTYSYIIIE-DKR 65
DB 15 GRFDIWAPEAGTVTLGGERYEMGRPGNGPADEGWTAADAPTADVDYGYLLDGEI 74

QY 66 EIPDPASRYQPLGVHDKSLQIRTDYQILDGKVKIEDL-----IYELHVGTFSSQ 115
DB 75 PLPDPTRRQPEGVH---ALSRT---FDPGAHRWQDAGWQRELOGSVIYELHIGTFTP 127

QY 116 EGNFKGVIEKLDYKDLGTIGTIELMPVAQPPGNRNGYDGVFLYAVONTYGGPWELAKLV 175
DB 128 EGTDLAAAGKLDYLAGLIGDIFIELLPVNAFNGTHNGYDGVQWFAVHEGYGPPAAQYQRFV 187

QY 176 NEAHKRGIAVILDVYVNHGTPGNYLLGLPYFSDRYKTPWGLTFNFDRCGDOVKFKIL 235
DB 188 DAAHAAGLVQDVVYVNHGTPGNYLLPRYGPYKLGEGNTWGSVNLDPGSDHVRQYIL 247

```
QY 236 ENVEYFKTFKIDGLRLDAVHAIFDNSPKHILQIEAKAHOLGK-----FVIAESDLND 289
:| | : : : : :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
Db 248 DNVAWMLRDYRVDGLRLDAVHALKDERAVHILEEFGALADALSSEGGRLPLTIAESDLNN 307
QY 290 PKIV--KDCGKIDQAWDDFHAVHAFITKEKDYIYQDFGRIDIEKTKDFVVDGK 347
:| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| |
Db 308 PRLLYPRDVNGYGLAGOWSDDFHVAHVNVSGETTGYYSDSLGALAKVLRDGGFFHDGS 367
QY 348 YSRVGRGTHGAPV--GDLPPRKFFVVFQNHQVGNRNGERLSILTDTKTYLMAATLYIL 405
:| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| |
Db 368 YSSFRGCHGRPINFSAVHPAALVVCSONHQIGNRATGDRLSOSLPYGSALAAVLTILT 427
QY 406 SPYIPLFMGEYYETNPFFFSDFSPVLIKGVREGRLKENNOM-----IDPQSEEA 458
:| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| |
Db 428 GPFTPMLFMGEYGATTPWQFFTSHPPELKGATAEGRIFERFMGWDPAVVPDQDPET 487
QY 459 FLKSKLSWK-----IDDEVLDYKQLNIRKRYNCKRVKVRREGNCITILMEKIGIIA 513
:| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| |
Db 488 FTRSKLDMAEASAGDHARLLELYSLITLR-----RSTPELARLGFADTAV----- 533
QY 514 SFDD 517
|||
Db 534 EFDD 537

RESULT 11
US-08-961-240-12
; Sequence 12, Application US/08961240
; Patent No. 5830715
; GENERAL INFORMATION:
; APPLICANT: KUBOTA, Michio
; APPLICANT: TSUSAKI, Keiji
; APPLICANT: HATTORI, Kazuko
; APPLICANT: SUGIMOTO, Toshiyuki
; TITLE OF INVENTION: DNA ENCODING ENZYME, RECOMBINANT DNA AND
; TRANSFORMANT, AND THEIR PREPARATIONS AND USE
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,240
; FILING DATE: 30-OCT-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/399,646
; FILING DATE: 07-MAR-1995
; APPLICATION NUMBER: JP 59834
; FILING DATE: 07-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 59840
; FILING DATE: 07-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: KUBOTA-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELE: 248633
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 596 amino acids
```

```
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-961-240-12

Query Match 32.8%; Score 969; DB 2; Length 596;
Best Local Similarity 39.5%; Pred. No. 3.2e-86;
Matches 215; Conservative 74; Mismatches 195; Indels 60; Gaps 12;

QY 13 GIFKLWAPYVNSVKLKLKSLIPMEKN-----DEGFEVEIDIEENLTSYIIIE-DKR 65
:| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| |
Db 15 GRFDIWAPEAGVTLLLAGGERYEMGRPGNGPADEGMMWTAADPTGADVDTGYLLDGDGEI 74
QY 66 EIDPPASRYOPLGVHDKSQLIRTDYQILDGLKVKIEDL-----IYELHWGTFSSQ 115
:| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| |
Db 75 PLDPRTTRQPEGVH---ALSRT---FDCAHWQDAGWQRELQSGVIYELHIGFTTP 127
QY 116 EGNPKGVIEKLDYLDKLGITGIELMPVAQFPNGRDWGVYGVLYAVONTYGGPWELAKLV 175
:| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| |
Db 128 EGTLDAAAGKLDYLAGLGIDFIELLPVNAFNGTHWGYDGVQWFAVHEGYGGAAYQRFV 187
QY 176 NEAHKRGIAVLDDVYNNHIGPEGNYLLGLGYPYFSDRYKTPWGLTFNEDDRCCDOVKFIL 235
:| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| |
Db 188 DAAHAAGLVQDVVYVNHLPSPGNYLPRYGPYKKGEGNTMGDSVNLDPGSDHVRQYIL 247
QY 236 ENVEYFKTFKIDGLRLDAVHAIFDNSPKHILQIEAKAHOLGK-----FVIAESDLND 289
:| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| |
Db 248 DNVAWMLRDYRVDGLRLDAVHALKDERAVHILEEFGALADALSSEGGRLPLTIAESDLNN 307
QY 290 PKIV--KDCGKIDQAWDDFHAVHAFITKEKDYIYQDFGRIDIEKTKDFVVDGK 347
:| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| |
Db 308 PRLLYPRDVNGYGLAGOWSDDFHVAHVNVSGETTGYYSDSLGALAKVLRDGGFFHDGS 367
QY 348 YSRVGRGTHGAPV--GDLPPRKFFVVFQNHQVGNRNGERLSILTDTKTYLMAATLYIL 405
:| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| |
Db 368 YSSFRGCHGRPINFSAVHPAALVVCSONHQIGNRATGDRLSOSLPYGSALAAVLTILT 427
QY 406 SPYIPLFMGEYYETNPFFFSDFSPVLIKGVREGRLKENNOM-----IDPQSEEA 458
:| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| |
Db 428 GPFTPMLFMGEYGATTPWQFFTSHPPELKGATAEGRIFERFMGWDPAVVPDQDPET 487
QY 459 FLKSKLSWK-----IDDEVLDYKQLNIRKRYNCKRVKVRREGNCITILMEKIGIIA 513
:| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| |
Db 488 FTRSKLDMAEASAGDHARLLELYSLITLR-----RSTPELARLGFADTAV----- 533
QY 514 SFDD 517
|||
Db 534 EFDD 537

RESULT 12
US-08-605-501-12
; Sequence 12, Application US/08605501
; Patent No. 5834287
; GENERAL INFORMATION:
; APPLICANT: KUBOTA, Michio
; APPLICANT: TSUSAKI, Keiji
; APPLICANT: HATTORI, Kazuko
; APPLICANT: SUGIMOTO, Toshiyuki
; TITLE OF INVENTION: DNA ENCODING ENZYME, RECOMBINANT DNA AND
; TRANSFORMANT, AND THEIR PREPARATIONS AND USE
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```


; FILING DATE: 07-MAR-1995
; APPLICATION NUMBER: JP 59834
; FILING DATE: 07-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 59840
; FILING DATE: 07-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: KUBOTA-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 597 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-961-240-4

Query Match 32.5%; Score 959; DB 2; Length 597;
 Best Local Similarity 40.2%; Pred. No. 3.1e-85;
 Matches 202; Conservative 81; Mismatches 186; Indels 34; Gaps 10;

QY	15	FKLWAPYVNSVKLKLKIPMEK-----NDGFFEVEIDDIIEENLTYSYIIE-DKREI	67
Db	18	YDWAFAESVTLLAGERYAMORRAETGPEDAGWTAGAPTDGNVDYGYLLDGDETPL	77
QY	68	PDPSRYQPLGVHDKSOLI-RTDYQILD---LGKVKIEDLIYIELHVGTFSQEGNFKGI	123
Db	78	PDPRTRQPDGVHLSKRTDPSAYSQDDDAWQGR-ELQGAIVYELHLGTFTEGLEAAA	136
QY	124	EKLIDYLDLIGITGIELMPVAQFPGNDWDGYDGVFLYAVQNTYGGPWELAKLYNEAHKGI	183
Db	137	GKLDYLAGLVDPFIELLPVNAFNGTINWGYDGVQWFAVHEAYGGPEAYQRFVDAHAAGL	196
QY	184	AVILDVVYVNHIGPEGNVLGLGYFSDRYKTPMGLTFNFDRCDOVRKFFILENVEYWFK	243
Db	197	GVLDQVYVNHGLPSGNYLPRFGPYLKQEGNTWGDSDVNLDPGSDHVRRIYLDNLAMWLR	256
QY	244	TFKIDGLRLDAVHAIFDNPSPKHILQIEIAEKHQLGKEV-----IAESDLNDPKIV--KD	295
Db	257	DVRVDGLRLDAVHALKDERAVHLEDFGALADQISAEVGRPLTLIAESDLNPNRLLYPRD	316
QY	296	DCGYKIDAQWDDFHAVHAFITKEKDYIQDFGRIEDIEKTKDVFYDGYDKYSRYRGRT	355
Db	317	VNGYGLEQWSDDFHAVHVNVTGETTYSDYDFSLAALAKVLRDGFHDGYSYSPREH	376
QY	356	HGAPV--GDLPPRKVFVFIQNHQVGNRNGERLSILTDTTYLMAATLYILSPIPLIF	413
Db	377	HGRPIFSAVHPAALVVCQNHQICGNRATGRLSOTLPYGSIALAAVLTGTFPTMLL	436
QY	414	MGEYETNPFFFDSPFLIKGVREGRLKENNOM-----IDPQSEEAFLKSKLSW	466
Db	437	MGEYGASTPWQFTSHPEPELGKATAEGRIKEFERMGWDPAVVPDPQDPETFRSKLDW	496
QY	467	KIDEE-----VLDYYKQLINIRK	484
Db	497	AEAAEGDHARLLELYRSLTALRR	519

Search completed: July 15, 2002, 12:18:05
Job time: 97 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 12:18:43 ; Search time 32.43 seconds
(without alignments)
1647.415 Million cell updates/sec

Title: US-09-298-924-8
Perfect score: 2954
Sequence: 1 MFSFGGNIENKKGIFKLWAP.....KLKDELKVNRGVGVQLE 556

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues
Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	2954	100.0	556	2 JC5132	alpha-amylase (EC
2	1784.5	60.4	561	2 S73087	alpha-amylase (EC
3	1748.5	59.2	558	2 JC5135	alpha-amylase (EC
4	1212.5	41.0	620	2 AH1827	alpha-amylase (imp
5	969	32.8	596	2 JC4697	trehalose trehaloh
6	959	32.5	598	2 S65770	malto-oligosyltreh
7	936	31.7	575	2 JC7727	probable hydrolase
8	900.5	30.5	594	2 AI0673	probable maltocollig
9	866.5	29.3	580	2 G70763	probable glycosyl
10	865.5	29.3	583	2 G83375	probable maltocollig
11	835.5	28.3	600	2 H75516	alpha-amylase Atu5
12	793.5	26.9	595	2 AG3194	probable 1,4-alpha
13	792.5	26.8	601	2 D96001	probable alpha-glu
14	493	16.7	552	2 S19134	hypothetical prote
15	493	16.7	552	2 AH1915	1,4-alpha-glucan b
16	353	11.9	666	2 B56639	glycogen operon pr
17	342.5	11.6	638	2 S18599	alkaline amylopull
18	333.5	11.3	707	2 S77094	alpha-amylase (EC
19	332.5	11.3	1280	2 E95031	pullulanase amyx
20	331.5	11.2	1256	2 G97902	1,4-alpha-glucan b
21	326	11.0	718	2 G69585	pullulanase amyx
22	322.5	10.9	422	2 S31839	1,4-alpha-glucan b
23	320	10.8	843	2 H72204	pullulanase - Ther
24	319.5	10.8	718	2 S73088	glycogen operon pr
25	315	10.7	562	2 S0312	alpha-amylase (EC
26	313.5	10.6	627	1 S40048	1,4-alpha-glucan b
27	311.5	10.5	642	2 E98000	1,4-alpha-glucan b
28	310.5	10.5	642	2 G95129	probable glycogen
29	306.5	10.4	666	2 E71565	

30	303.5	10.3	664	2 B86539	glycogen hydrolase
31	303.5	10.3	664	2 B72084	glycosyl hydrolase
32	302	10.2	783	2 B84823	probable isoamylas
33	296.5	10.0	764	2 S47569	1,4-alpha-glucan b
34	296.5	10.0	818	2 T02231	probable isoamylas
35	296	10.0	630	2 D70363	1,4-alpha-glucan b
36	295	10.0	666	2 G81717	glycosyl hydrolase
37	295	10.0	789	2 T01321	probable isoamylas
38	293.5	9.9	717	2 H84057	pullulanase amyx l
39	292.5	9.9	733	2 T49679	probable branching
40	292	9.9	1072	2 A84112	alkaline amylopull
41	289	9.8	720	2 G97229	pullulanase (impor
42	289	9.8	741	1 S34218	1,4-alpha-glucan b
43	286.5	9.7	666	2 D82511	1,4-alpha-glucan b
44	283	9.6	746	2 S74546	glycogen operon pr
45	282	9.5	770	2 S76095	hypothetical prote

ALIGNMENTS

RESULT 1

JC5132

alpha-amylase (EC 3.2.1.1) - Sulfolobus acidocaldarius

N;Alternate names: glycogenase

C:Species: Sulfolobus acidocaldarius

C>Date: 31-Jan-1997 #sequence_revision 27-Feb-1997 #text_change 20-Jun-2000

C:Accession: JC5132

R; Kobayashi, K.; Kato, M.; Miura, Y.; Kettoku, M.; Kameda, T.; Iwamatsu, A.

Biosci. Biotechnol. Biochem. 60, 1720-1723, 1996

A; Title: Gene analysis of trehalose-producing enzymes from hyperthermophilic archaea

A; Reference number: JC5131; MUID:97141330

A; Accession: JC5132

A; Status: preliminary; nucleic acid sequence not shown

A; Molecule type: DNA

A; Residues: 1-556 <KOB>

A; Cross-references: DDBJ: D64131; NID: g987704; PIDN: BAA11011.1; PID: g987705

A; Experimental source: strain ATCC33909

C; Comment: This enzyme hydrolyzes glycosyltrehalose to liberate trehalose.

C; Superfamily: trehalose trehalohydrolase

C; Keywords: glycosidase; hydrolase

Query Match 100.0%; Score 2954; DB 2; Length 556;
Best Local Similarity 100.0%; Pred No. 2,3e-194;
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFSFGGNIENKKGIFKLWAPYVNSVKLSKLIIPMEKNDGEGFEVEIDDIENLTYSYI 60

Db 1 MFSFGGNIENKKGIFKLWAPYVNSVKLSKLIIPMEKNDGEGFEVEIDDIENLTYSYI 60

QY 61 IEDKREIPDPASRYQPLGVHDKSQLIRTDYQILDGKVKIEDLIYELHVGTFSGGNPK 120

Db 61 IEDKREIPDPASRYQPLGVHDKSQLIRTDYQILDGKVKIEDLIYELHVGTFSGGNPK 120

QY 121 GVTEKLDYLDLKITGTELMPVAQFPGNRDWDGVDGVFLAVQNTYGGPWEAKLVNEAHK 180

Db 121 GVTEKLDYLDLKITGTELMPVAQFPGNRDWDGVDGVFLAVQNTYGGPWEAKLVNEAHK 180

QY 181 RGIAVILDVYVNHIGPEGNYLLGLGPFYSDRYKTPNGLTGTFNFDGRCQDQVRKFIENVEY 240

Db 181 RGIAVILDVYVNHIGPEGNYLLGLGPFYSDRYKTPNGLTGTFNFDGRCQDQVRKFIENVEY 240

QY 241 WFKTFKIDGLRLDAVIAFDNSPKHILOETAEKAHQLGFVIAESDLNDPKIYVDDCGYK 300

Db 241 WFKTFKIDGLRLDAVIAFDNSPKHILOETAEKAHQLGFVIAESDLNDPKIYVDDCGYK 300

QY 301 IDAQWDDDFHHAVHAFITTEKDYVYODFGRIEDIETKFDVYDGYKSYRGRTHGAPV 360

Db 301 IDAQWDDDFHHAVHAFITTEKDYVYODFGRIEDIETKFDVYDGYKSYRGRTHGAPV 360

QY 361 GDLPFRKRVVFIQNHQVQNGRNGERLSILTDTKTTYIMATLYILSPYIPLIFMGEEYE 420

|||||

Db 361 GDLPRKVFVFIQNHQVGNRNGERISILTDKTTYLMAATLYILSPYIPLIFMGEEYYE 420
Qy 421 TNPEFFSDFSDPVLKGVREGRLKNNQMDPQSEAFKSKLSWKIDIEVDLYYKOLI 480
Db 421 TNPEFFSDFSDPVLKGVREGRLKNNQMDPQSEAFKSKLSWKIDIEVDLYYKOLI 480
Qy 481 NIKRYNCKRVKVRREGNCITLIMEKIGIIASFDIVINSKITGNLLIGIGPKPKLKK 540
Db 481 NIKRYNCKRVKVRREGNCITLIMEKIGIIASFDIVINSKITGNLLIGIGPKPKLKK 540
Qy 541 DELIKVNRGVGVYOLE 556
Db 541 DELIKVNRGVGVYOLE 556
RESULT 2
S73087
alpha-amylase (EC 3.2.1.1) precursor - Sulfolobus solfataricus
N:Alternate names: protein c0620
C:Species: Sulfolobus solfataricus
C:Date: 09-Oct-1997 #sequence_revision 24-Oct-1997 #text_change 20-Jun-2000
C:Accession: S73087
R:Sensen, C.W.; Klenk, H.P.; Singh, R.K.; Allard, G.; Chan, C.C.Y.; Liu, Q.Y.; Penny, S.
Mol. Microbiol. 22, 175-191, 1996
A:Title: Organizational characteristics and information content of an archaeal genome: I
A:Reference number: S73076; MUID:97055432
A:Accession: S73087
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-561 <SEN>
A:Cross-references: EMBL:Y08256; NID:gl707679; PIDN:CAA69503.1; PID:gl707699
A:Experimental source: strain p2
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1996
C:Superfamily: trehalose trehalohydrolase
C:Keywords: glycosidase; hydrolase
Query Match 60.4%; Score 1784.5; DB 2; Length 561;
Best Local Similarity 59.2%; Pred. No. 2.4e-114;
Matches 332; Conservative 90; Mismatches 130; Indels 9; Gaps 6;
Qy 3 SFGNIEKNKGIFKLNAPYVNSVKL--LSKKLIPMEKNDEGFEVEIDDEENLTSYII 61
Db 2 TFGYKLEDGVTNFWAPYQKVKLTNRGIEMERDDKGYFTITLNNVGRDVKYIL 61
Qy 62 EDKREIPDPASRQPLGVHDKSOLIRTDYQILDGLGVKIEDLIYELHWGTFSGBNF 119
Db 62 DDNSEVPDPASRQPEGVHGYSEIISPDFWDDNSKVKREDLVIYELHIGFTSEGT 121
Qy 120 KGVTEKLDYKDLGITGIELMPVAQPPGNRDWGYDGVFLYAVONTYGGPWELAKLVNEAH 179
Db 122 EGVTKLNLKELGVTAIEIPIAQPFGKDWGYDGVLYAVONSYYGSGPGRKLVNEAH 181
Qy 180 KRGTAVIDVYVNHGPEGNYLLGLGYPFSDRYKTPWGLTFNFDRCDOVRFEILENVE 239
Db 182 KGLAVILDVYVNHVGPENYMKLGYPFSEKTKPWGLTFNFDAGSDEVRFEILENVE 241
Qy 240 YWFKTKIDGLDLDAVHAIFDNPSPKHILQIEIAEKHQLGKFIASDLNDPKIV--KDDC 297
Db 242 YWNEFVHVDGFLDAVHAITDNPSPKHILEADIADVHKYDIAVIAESDLNDPRVVPKEK 301
Qy 298 GYKIDAQWDDPHAVHAFITKEKDYIYODFGRIEDIEKTFKDVYVYDGYKSYRGRTHG 357
Db 302 GYNIDAQWDDFHAHAFTLGERQGYSDFGSGIDGVKSYKDVFIYDGYKSNFRKTHG 361
Qy 358 APVGDLPFRFVFIQNHQVGNRNGERISILTDKTTYLMAATLYILSPYIPLIFMGEE 417
Db 362 KSVGDGDCRFVYVFIQNHQVGNRNGERILKLVKDESKYKIAAALYILSPYIPIFMGEE 421
Qy 418 YETNPFFSDFSDPVLKGVREGRLKNNQMDPQSEAFKSKLSWKIDIEVDLYYK 477
Db 422 YGENPYFYSDPKPLKGVREGRRNGEQETDPOSDCTFNDKSLKSWINDDILSYFK 481

Qy 478 QLNIKRYN--NCKRVKVRREGNCITLIMEKIGIIASFDIVINSKITGNLLI--GIGF 534
Db 482 SLKIRKEYGLACNKLKSVENGNTWLVKNGCLAVVVFVSKSVIEMKYSGLTVLSSNSF 541
Qy 535 PKLKLKDELKLVNRGVGVYOL 555
Db 542 PSQITESH--YELDKGFALYKL 561
RESULT 3
JC5135
alpha-amylase (EC 3.2.1.1) - Sulfolobus solfataricus
C:Species: Sulfolobus solfataricus
C:Date: 31-Jan-1997 #sequence_revision 27-Feb-1997 #text_change 20-Jun-2000
C:Accession: JC5135
R:Kobayashi, K.; Kato, M.; Miura, Y.; Kettoku, M.; Kameda, T.; Iwanatsu, A.
Biosci. Biotechnol. Biochem. 60, 1882-1885, 1996
A:Title: Gene cloning and expression of new trehalose-producing enzymes from the hype
A:Reference number: JC5134; MUID:97141610
A:Accession: JC5135
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-558 <KOB>
A:Cross-references: DDBJ:D641130; NID:g987702; PIDN:BA11010.1; PID:gl838936
A:Experimental source: strain KMI
C:Comment: This enzyme hydrolyzes starch alpha-1,4 linkage adjacent to the alpha-1,1
C:Superfamily: trehalose trehalohydrolase
C:Keywords: glycosidase; hydrolase
Query Match 59.2%; Score 1748.5; DB 2; Length 558;
Best Local Similarity 58.9%; Pred. No. 6.9e-112;
Matches 329; Conservative 87; Mismatches 136; Indels 7; Gaps 5;
Qy 3 SFGNIEKNKGIFKLNAPYVNSVKL--LSKKLIPMEKNDEGFEVEIDDEENLTSYII 61
Db 1 TFAFKIDGNEVIFTWAPYQKVKLVLEKGLYEMERDEKGYFTITLNNVGRDVKYVL 60
Qy 62 EDKREIPDPASRQPLGVHDKSOLIRTDYQILDGLGVKIEDLIYELHWGTFSGBNF 121
Db 61 DDASEVPDPASRQPEGVHGYSEIISPDFWDDNSKVKREDLVIYELHIGFTSEGT 120
Qy 122 VIEKLDYKDLGITGIELMPVAQPPGNRDWGYDGVFLYAVONTYGGPWELAKLVNEAH 181
Db 121 VIEKLDYKDLGITGIELMPVAQPPGNRDWGYDGVFLYAVONSYYGSGPGRKLVDEAHK 180
Qy 182 GIAVILDVYVNHGPEGNYLLGLGYPFSDRYKTPWGLTFNFDRCDOVRFEILENVE 241
Db 181 GLGVILDVYVNHVGPENYMKLGYPFSEKTKPWGLTFNFDAGSDEVRFEILENVE 240
Qy 242 FKTFKIDGLDLDAVHAIFDNPSPKHILQIEIAEKHQLGKFIASDLNDPKIV--KDDCY 299
Db 241 IKEYNDGFLDAVHAITDNPSPKHILEADIADVHKYDIAVIAESDLNDPRVVPKEK 300
Qy 300 KIDAQWDDPHAVHAFITKEKDYIYODFGRIEDIEKTFKDVYVYDGYKSYRGRTHG 359
Db 301 NIDAQWDDPHHSHAYLTGERQGYTDFGNLDDIVKSYKDVYDGYKSYNFRKTHG 360
Qy 360 VGDLPFRKVFVFIQNHQVGNRNGERISILTDKTTYLMAATLYILSPYIPLIFMGEE 419
Db 361 VGELDCNFFVFIQNHQVGNRNGERITKLVDRESKYIAAALYILSPYIPIFMGEE 420
Qy 420 ETNPFFSDFSDPVLKGVREGRLKNNQMDPQSEAFKSKLSWKIDIEVDLYYKOL 479
Db 421 EENPFFSDFSDSKLIQVREGRRKENGQDTPQDESTENASKLSWKIDIEFSEFKIL 480
Qy 480 INIKRYN--NCKRVKVRREGNCITLIMEKIGIIASFDIVINSKITGNLLIGI--GFPK 536
Db 481 IKMKELSIACDRRVNVNNGENWLIITKGREYSLYVFSKSYIEVYSGTLLSSNSFPQ 540
Qy 537 KKLKDELKLVNRGVGVYOL 555
Db 541 HIEEGK--YEPDKGFALYKL 558

A:Molecule type: DNA
A:Residues: 1-598 <MAR>
A:Cross-references: EMBL:D63343; NID:g1255444; PIDN:BAA09668.1; PID:g1255446
C:Superfamily: trehalose trehalohydrolase

Query Match 32.5%; Score 959; DB 2; Length 598;
Best Local Similarity 40.2%; Pred. No. 7.8e-58;
Matches 202; Conservative 81; Mismatches 186; Indels 34; Gaps 10;

Qy 15 FKLWAPYVNSVKLKLIPMEK-----NDEGFEVEIDIEENLTYSYIIIE-DKREI 67
Db 19 YDWWAPNAESVTLLAGGERYAMQRAETGPEADGAWMTAAGAPTNGVDVGYLLDGETPL 78
Qy 68 PDPAASRQPLGVHDKSQLI-RTDQIILD--LGKVKIEDLIYELHVGTFSGQEGNFKGVI 123
Db 79 PDPRTRQPCGVHSLRTFDSAYSWQDDAQGR-ELQGAIVYELHGLTFTPEGTLEAA 137
Qy 124 EKLDYLDGLGTIGTIELMPVAFPGNRMGVDGVFLYAVQNTYGGPWELAKLVNEAHKRG 183
Db 138 GKLDYLAGLVDFTELLPVNAFNGTHNWGYDGVQVFAVEAYGGPEAYQRFVDAHAAGL 197
Qy 184 AVILDVYVNHGTPGNYLLGLGPFSDRYKTPWGLTFNFDGRCDOYRKFFILENVVWF 243
Db 198 GVIQDVVYVNHGTPGNYLLPREGPYLYKOGEGNTWGDVNLDPGSDHVRYYILDLNML 257
Qy 244 TFKIDGLRLDAVHAIFDINSRKHILOEIAEKAKHQLGKFV-----IAESDLNDPKIV--KD 295
Db 258 DYRVDGLRLDAVHALKDERAVHILEDFCALADQISAEVGRPLTLIAESDLNPRLLYPRD 317
Qy 296 DCGYKIDAQWDDPHHVAHFATIEKDYKDYODFGRIEDIEKTKDFVYDGKYSRYRGRT 355
Db 318 VNGYGLEQWSDDFHVAHVNVGTETGYYSDFDLSAALAKVLRDGFHDCGYSYFSFRERH 377
Qy 356 HGAPV--GDLPRFVVFQIHNHDOVNGRNGERLSILTOKTLYLMAATLYLTSPYIPLIF 413
Db 378 HGRFINSAVHPAALVCSQNHQDIQGNRATGDRLSQTLPGSLAALAVLTUTGFTFMILL 437
Qy 414 MGEYETNPFFSDFSDPVLKIGVREGRLKENNQ-----IDPQSEAFKLSLW 466
Db 438 MGEYGASTPWQFTSHPELPGKATAEGRIKEFERMGWDPAVVPDQDPETFRSKLDW 497
Qy 467 KIDEE-----VLDYKQLINIRK 484
Db 498 AEAAGDGHARLLELYRSILALRR 520

RESULT 7
JC7727
malto-oligosyltrehalose trehalohydrolase (EC 3.2.1.141) - Arthrobacter ramosus
C:Species: Arthrobacter ramosus
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 31-Dec-2001
C:Accession: JC7727; PC7175
R:Yamamoto, T.; Maruta, K.; Watanabe, H.; Yamashita, H.; Kubota, M.; Fukuda, S.; Kurimoto
Biosci. Biotechnol. Biochem. 65, 1419-1423, 2001
A:Title: Trehalose-producing operon trex2 from Arthrobacter ramosus S34.
A:Reference number: JC7726; MUID:21364352; PMID:11471747
A:Accession: JC7727
A:Molecule type: DNA
A:Residues: 1-575 <YAM>
A:Cross-references: DDBJ:AB045141
A:Experimental source: strain S34
A:Accession: PC7175
A:Molecule type: protein
A:Residues: 1-20;31-50;298-319 <YA2>
C:Comment: This enzyme is involved in producing trehalose from glycogen.

C:Genetics:
A:Gene: tre2
C:Superfamily: trehalose trehalohydrolase
C:keywords: glycosidase; hydrolase

Query Match 31.7%; Score 936; DB 2; Length 575;

Best Local Similarity 38.7%; Pred. No. 2.8e-56;
Matches 198; Conservative 82; Mismatches 177; Indels 54; Gaps 10;

Qy 15 FKLWAPYVNSVKLKL--SKKLIPMEKNDGFFVE--IDDIEENLTYSYIIIEKREIPDP 70
Db 5 FPWAPQAAQVTLVVGGRAEPLTRDENGWALQQPWDGGPDLVDYGLVGGKPPFADP 64
Qy 71 ASRYQPLGVHDKSQLIRTDYQIILDGKV-----KIEDLIYELHVGTF 114
Db 65 RSLRQPRGVH-----ELGREFDPARYANGDDGWRGRDLTGAVIYELHVGTF 111
Qy 115 QEGNFKGVIEKLDYLDGLGTIGTIELMPVAFPGNRMGVDGVFLYAVQNTYGGPWELAKL 174
Db 112 PEGTLDIAIRRLDLHLVGLGVDAVELLPVNAFNGTHWGYDGVLMVAVHEPYGGPEAYQRF 171
Qy 175 VNEAHKRGIAVLDDVYVNHGTPGNYLLGLGPFSDRYKTPWGLTFNFDGRCDOYRKFI 234
Db 172 VDACHAGLVAQVQVYVYVNHGTPGNGNHLDPFGPYLGSGAANTWGDALNDGPLSDEVRYI 231
Qy 235 LENVEYFKTFKIDGLRLDAVHAIFDINSRKHILOEIA-----EKAHQLGK--FVIAESDLN 288
Db 232 IDNAVYWLROMHADGLRLDAVHALDARALHLELAARVDELAGEGRPLTLIAESDLN 291
Qy 289 DPKIVKDDC--GYKIDAQWDDPHHVAHFATIEKDYKDYODFGRIEDIEKTKDFVYVDG 346
Db 292 DPKLIRSAAHGYGLDAQWDDVHHAVHNVGTETGYVYADFGLGALVKVFORGFWDG 351
Qy 347 KYSRYRGRTHGAPVG-DLPPRKFFVFQIHNHDOVNGRNGERLSILTOKTLYLMAATLYL 405
Db 352 TWSSFRHHRGLDPLDIPFRRLVAFADQDDQGNRAVGRMSAQVGEGLSAAAAALLVLL 411
Qy 406 SPYIPLFMGEYETNPFFSDFSDPVLKIGVREGRLKENNQ-----IDPQSEEA 458
Db 412 GPFTPLFMGEWCAETPWQFTSHPELGEATARGIAEFARMGWDPAVVPDQDPAT 471
Qy 459 FLASKLSWKIDEE-----VLDYKQLINIRK 484
Db 472 FARSHLDWSEPEREPHAGLLAFATYDILALRR 502

RESULT 8
AI0673
probable hydrolase STY1503 [imported] - Salmonella enterica subsp. enterica serovar T
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: this species has also been called Salmonella typhi
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 27-Nov-2001
C:Accession: AI0673
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farr
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens,
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica se
A:Reference number: AB0502; PMID:11677608
A:Accession: AI0673
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-594 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD01762.1; PID:g16502610; GSPDB:GN001776
C:Genetics:
A:Gene: STY1503
C:Superfamily: trehalose trehalohydrolase

Query Match 30.5%; Score 900.5; DB 2; Length 594;
Best Local Similarity 38.9%; Pred. No. 7.8e-54;
Matches 199; Conservative 79; Mismatches 173; Indels 61; Gaps 13;

Qy 15 FKLWAPYVNSVKLKLIPMEKNDGFFVEIDIEENLTYSYIIIEKREIPDPASRY 74
Db 22 FRLWATGQKVMRLACKDQEMQASGSGWFTLDVSGVTPGTETYNFLVSDGMVLPDPASRA 81
Qy 75 QPLGVHDKSQLIRTDYQIILDGKV-----KIEDLIYELHVGTFSGEGNFKGVIE 124

Db 82 QKTDVNGPS-----YVIDPGSYAWRNTGWKSRWEQAVVYEMHTGTTPGTFHTAIA 134
QY 125 KLDYLKDLGITGIELMPVAFQPCNDRWGDYGVFLYAVONTYGGPWELAKLVNEAHRKRTIA 184
Db 135 KLPYLAELGVTVIEWPVAQFGGERGWGDYGVLLYAPHSAYGTPDDFAFIDTAHGYGLS 194
QY 185 VILDVYVNHIGPEGNYLLGLSP-YFSDRYKTPWGLTFNFDRCGQOVRKFIENVEYWFK 243
Db 195 VILDVNLHFGPEGNYLLPLAFPHKRWTPWNGIAYD---VDAVRYIITEAPLYWLT 251
QY 244 TKIDGLRLDAVHAFDNPSPKILQIEAKAHQLGKFVIAESDLNDPKI----- 292
Db 252 EYHLDGLRFDADIQIEDSSARHVLVEIAQRIE-----DITDRPLHLTJEDSRNII 302
QY 293 ---VKDDCYK-IDAQWDDHFAHAFITKEKDYDYDFGRIED-IEKTFKDVYVD 345
Db 303 SLHPRDODGNAPLFAEWNDDEHNAVHVFATGETQYNYNDFADAPKHLARALAEAFYQ 362
QY 346 GKYSRYGRTHGAPVGDLPKPKFVFIQNHQVGNRNGERLSILTD-KTTYLMAATLYI 404
Db 363 GEISPOTGPRGVKSGQPPVAFVDFIQNHQVGNRAQGRDLITLAGAERTKVLLATL-L 421
QY 405 LSPYIPLPMGEYETNFFFSDFSDPVLKGVREGRLKE-----NNQMDPQSEAF 459
Db 422 LSPHILPLPMGEYGESRPFLFTDFHGD-LARAVREGRAKEFADHAGENVPDPNAPETF 480
QY 460 LKSKLSWKIDEEV-----LDYKQLINIRKY 486
Db 481 QRSKLNWQKHSEKAWLAFTRELLLLKQKH 512

RESULT 9
G70763
probable maltoligosyltrehalose trehalohydrolase - Mycobacterium tuberculosis (strain H37Rv)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C:Accession: G70763
R:Colo, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987
C:Accession: G70763
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-380 <COL>
A:Cross-references: GB:274020; GB:AL123456; NID:g3261584; PIDN:CRA98329.1; PID:g1403480
A:Experimental source: strain H37Rv
C:Genetics:
A:Gene: glqZ
C:Superfamily: trehalose trehalohydrolase

Query Match 29.3%; Score 866.5; DB 2; Length 580;
Best Local Similarity 37.3%; Pred. No. 1.6e-51;
Matches 187; Conservative 83; Mismatches 194; Indels 37; Gaps 11;
QY 15 FKLWAPYVNSVKLSKLIPEKNDGFEVEIDIEENLYSYIIEKREI-PDPASR 73
Db 4 FRWAPKPAVLRLVDNGAVHAMTRSDGWWHTV-AAPADARYGYLLDDDPVLPDPRSA 62
QY 74 YQPLGVHDKSQ-----LLRTDYQILDGLGVKIEDLIYELHVGTFSQSGNFKGVLE 124
Db 63 RQPDGVHARSQRWPEPPQFGAARTD---TGWPGRSVEGAVIYELHIGTETTAGTFDAALIE 119
QY 125 KLDYLKDLGITGIELMPVAFQPCNDRWGDYGVFLYAVONTYGGPWELAKLVNEAHRKRTIA 184
Db 120 KLDYLVLDIGIDVELMPVNSFAGTGWGDYGVLYSVHPEYGGPDGLRFIDACHARRLG 179
QY 185 VILDVYVNHIGPEGNYLLGLSPYFSDRYKTPWGLTFNFDRCGQOVRKFIENVEYWFK 244

Db 180 VLIDAVFNHLGFSNGYLPFRGPLYLSSA-SNPWGDGINIAGADSDEVRYHIIDCALRWMD 238
QY 245 FKIDGLRLDAVHAFIDNPSPKILQIEAKAH---OLGK-FVIAESDLNDPKIV--KDD 296
Db 239 FIADGLRLDAVHALVDVTTAVHLEELANATRWLSQGLRPLSLIAETDRNDPLITRPSH 298
QY 297 CGYKIDQAVDDHFAHAFITKEKDYDYDFGRIEDIEKTFKDVYDGKYSRYGRTH 356
Db 299 GGYGITAQWDDHFAHAFITAVSGERGQYVADFGSLATLAYTLRNGYFHAGTYSRFRHRH 358
QY 357 GAPV--CDLPKPKFVFIQNHQVGNRNGERLSILTDKTTYLMAATLYLSPYIPLIM 414
Db 359 GRALDTSATPATRLLEYTCTHDPQVGNRALGDRPSQYLTGGQLAKAALTIGSPYATAMLF 418
QY 415 GREYETNFFFSDFSDPVLKGVREGRLKE-----NNQMDPQSEAFKLSKSNK 467
Db 419 GEWGASSPFCFCSHPELAHSIVAGRKEFAEHGWAADDIPDPQDPQTFQRCKLNWA 478
QY 468 -----IDEEVLDYKOLINIR 483
Db 479 EAGSGEHAHLRHFYRDLIALR 499
RESULT 10
G83375
probable glycosyl hydrolase PA2164 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: G83375
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa
A:Reference number: A82950; MUID:20437337
C:Accession: G83375
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-583 <STO>
A:Cross-references: GB:AE004643; GB:AE004091; NID:g9948178; PIDN:AAG05552.1; GSPDB:G
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA2164
C:Superfamily: trehalose trehalohydrolase

Query Match 29.3%; Score 865.5; DB 2; Length 583;
Best Local Similarity 38.2%; Pred. No. 1.9e-51;
Matches 199; Conservative 89; Mismatches 166; Indels 67; Gaps 20;
QY 4 FGNIEKN-KGTFKLWAPYVNSVKLSK-KLIPMEKNDGFEVEIDIEENLYSYII 61
Db 7 FGAQFGCGNRTCFGLWAPDAREVRVETADGRDWPLEGSGWFEATL-PCPGTRTRYRI 65
QY 62 EDKREIPDASRYQPLGVHDKSQLIRTDYQILDGLKV--KIE-----DLIYELHVG 111
Db 66 DGRPGVPDPASQFPQDGVHGS-----QVLHDGYANRWDEWRGPRWHEAVIYELHVG 118
QY 112 TFSQENFGKVIKLDYLDLKITGIELMPVAFQPCNDRWGDYGVFLYAVONTYGGPWEL 171
Db 119 LF---GSYAEVERFLPRLVELGVTAVELMPLGFEFGRRNMGYDGLVLPFAPASAYGTPQL 175
QY 172 AKLVNEAHRKRTIAVILDVYVNHIGPEGNYLLG-LGPFYSDRYKTPWGLTFNFDRCGQV 230
Db 176 KHLIDSAHGMGLMVFVDYIYNHFGPCNLAQYAAAFRRDRQTPWGQALDF-RRG--EV 232
QY 231 RKFIENVEYWFKTFKIDGLRLDAVHAFIDNPSPKILQIEAKAH-----QLGKFVIAB 285
Db 233 REFFYNALMWLLDYRVDGLRFDVHAIPDSA---FLVEMARRLRGAAGPERHVLHVLN 289
QY 286 DUNDPKIVKDDCGYKIDAQWDDFHAVHAFITKEKDYDYQDFGR-IEDIEKTFKDVFY 344

Db 290 DNRASLLRQ--GY--DAQWDDGHHALHVLITGENDGYQDYPPBLRCLARCLAEFFVY 345

QY 345 DQKYSRYRGRTUGAPVGLDPPKFFVFIQNHQVQGNRNGERLSILTDKTTYLMAATLYI 404

Db 346 QGEANRH-GRPRGEPADLAPDAFVLFQNHQVQGNRNGERLSVLAEPQALRLAIALQL 404

QY 405 LSPYIPLIFMGEEYETNPFPPFFSDPVLKIKVREGRLKENNQ-----MID 452

Db 405 LAPMPLIFMGEECAARPFLLFTDHQGE-LADAVREGRRKEFGFGRFGEGATLASLPD 463

QY 453 POSEBAFLKSK-----LSWKIDEVLDYKOLINIRKRY 486

Db 464 PNAVETFRSRLACDPAWR-----GFRQLLEIRHEH 498

RESULT 11

H75516

Probable maltotigosyltrehalose trehalohydrolase - Deinococcus radiodurans (strain R1)

C:Species: Deinococcus radiodurans

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000

C:Accession: H75516

R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A:Reference number: A75250; MUID:20036896

A:Accession: H75516

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-600 <WHI>

A:Cross-references: GB:AE001905; GB:AE000513; NID:g6458144; PIDN:AAF10042.1; PID:g645814

C:Genetics:

A:Map position: 1

C:Superfamily: trehalose trehalohydrolase

Query Match 28.3%; Score 835.5; DB 2; Length 600;

Best Local Similarity 36.8%; Pred. No. 2.2e-49;

Matches 181; Conservative 89; Mismatches 191; Indels 31; Gaps 12;

QY 15 FKLWAPYVNSVKLKLKSLIPMEKNDGFEVEIDDEENLYTYIIEIDKREIPDPASRY 74

Db 38 FLMTSTARTVAVRVNGTEHVMTSLGGGIYELEL-PVGPARYLFVL-DGVPTDPDYARF 95

QY 75 QPLGVHDKSQLI---RTDYQILDGLCKVKIEDLIYELHVGTFSGEKNFGVIEKLDYLKD 131

Db 96 LPDGVHGEAEVVDGTFDWDADWHGKGLADCVFEYEVHVGTFTEGYYRAAAEKLPYLKE 155

QY 132 LGITGIELMPVAOPGPNRDMGVDGVFLYAVQNTYGGPWELAKLYNEAHKRGIAVILDVY 191

Db 156 LGVTAIQWMLAAPDGGQWGYDGAFAVPYGRPEDLMALVDAAHRLGLGVFLDVY 215

QY 192 NHIGPEGNYLLGLCP-YFSDRYKTPWGLTFNFDRCQDQVKKFLENVVFWKTFKIDGL 250

Db 216 NHFGPSLLSSYAPSYFTDRFSAWGMGLDYAE---PHMRRYVTGNARWMLRDYHFDGL 272

QY 251 RLDAVHAIFNSPKHILQETAETHAQLG---KFVIAESDLNDPKIVKDCGKIDAQWDD 308

Db 273 RLDAATPYMTDSETHILTELAQETHELGCTHLLAEHRNLPDLVTYN---HLDGIWTD 329

QY 309 FHIAVHAFTKEKDYQDF-GRIEDIKTFKDVYVYDGKYSYRGTRH--GAPVGDLP 365

Db 330 FHHETRVTLTGEGYVAGYRGGAELAYTIRGRWRYEGQFWAVKGEERHGPSDALEA 389

QY 366 RKFWVFQNHQVQGNRNGERLSILTDKTTYLMAATLYIILSPYIPLIFMGEEYETNPF 425

Db 390 PNFWVFQNHQVQGNRNGERLSILTDKTTYLMAATLYIILSPYIPLIFMGEEYETNPF 449

QY 426 FFDSDPVLKIKVREGRLKE-----NNQIMDPQSEAEFLKSLKSWIDE-----EV 472

Db 450 FFSDHAGE-LGOAVSEGRKKFEGFSGFSGEDVDPDQAEQTFLSNKLNAWREGEHART 508

QY 473 LDYKOLINIRK 484

Db 509 LRLYRDLLRLR 520

RESULT 12

AG3194

alpha-amylase Atu5284 [imported] - Agrobacterium tumefaciens (strain C58, Dupont) pla

C:Species: Agrobacterium tumefaciens

C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 01-Feb-2002

C:Accession: AG3194

R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McCI : Karp, P.; Romero, P.; Zhang, S.

Science 294, 2317-2333, 2001

A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kam ster, E.W.

A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A:Reference number: AB2577; PMID:11743193

A:Accession: AG3194

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-595 <KUR>

A:Cross-references: GB:AE008687; PIDN:AAL45973.1; PID:g17743726; GSPDB:GN00188

A:Experimental source: strain C58 (Dupont)

C:Genetics:

A:Gene: Atu5284

A:Genome: plasmid

C:Superfamily: trehalose trehalohydrolase

Query Match 26.9%; Score 793.5; DB 2; Length 595;

Best Local Similarity 36.7%; Pred. No. 1.6e-46;

Matches 175; Conservative 76; Mismatches 195; Indels 31; Gaps 11;

QY 15 FKLWAPYVNSVKLKLKSLIPMEKNDGFEVEIDDEENLYTYIIEIDKREIPDPASRY 74

Db 22 FSIWAPASATVKLMLNDAEFDMHTAGDGHDTKPALPGD-RYGFVLADTRFVADPASNR 80

QY 75 QPLGVHDKSQLIRTDYQILD---LGKVKIEDLIYELHVGTFSGEKNFGVIEKLDYLKD 131

Db 81 QQEGPRGPSLIVNHDFAWKPNWKR-PWHEAVVYELHIGTFTEGTFAAAAEKLEVLAD 139

QY 132 LGITGIELMPVAOPGPNRDMGVDGVFLYAVQNTYGGPWELAKLYNEAHKRGIAVILDVY 191

Db 140 VGITIELMPLATFAGSRGWGIDGVLOFSPORDYGTDELKAFIDQAHGHGIMVLDVY 199

QY 192 NHIGPEGNYLLGLGPFYSDRYKTPWGLTFNFDRCQDQVKKFLENVVFWKTFKIDGLR 251

Db 200 NHFGPAGNTLQAYAPAFKKHETPWGPAPDENRA---EVSFFFLQAFYWLTYRFDGLR 256

QY 252 LDV-HAIFNSPKHILQETAETHAQLG----FVIAESDLN--DPKIVDCCGYKIDAQ 304

Db 257 IDAADHLAGDGEVDFLIEMAREVKRTIRNRHVLHVEDARNAAASPMTPMADGAILVDAQ 316

QY 305 WVDHFHVAHFITKEKDYIYQDFR--IEDIEKTFKDVYVYDGKYSYRG-RTHGAPVG 361

Db 317 WNDHFHVIHVATNNEGGIYEDFASPYENLRSLATGTYQGEPRPSRNFASGEP 376

QY 362 DLPPKRVFVFIQNHQVQGNRNGERLSILTDKTTYLMAATLYIILSPYIPLIFMGEEYET 421

Db 377 HLPPHREVFNLHNDQAGNRLGERLALPPPLFGTLEALLCPTPLVFMGDEHGA 436

QY 422 NPFFSDFSDPVLKIKVREGRLKE-----NNQIMDPQSEAEFLKSLW 466

Db 437 NPFFSDFSDPHNREQEIR-NRLKAESFQGEPLPPDASQVMVMDPDQHTMQLTSLKW 492

RESULT 13

D96001

Probable 1,4-alpha-glucan branching enzyme (EC 2.4.1.18) [imported] - Sinorhizobium m

C.Species: Sinorhizobium meliloti
C.Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 14-Sep-2001
C.Accession: D96001
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A.Title: The complete sequence of the 1,683-kb pSymb megaplasmid from the N2-fixing endo
A.Reference number: A95842; MUID:21396508; PMID:11481431
A.Accession: D96001
A.Status: preliminary
A.Molecule type: DNA
A.Residues: 1-601 <KUR>
A.Cross-references: GB:AL591985; PIDN:CAC49676.1; PID:gl5141163; GSPDB:GN00167
A.Experimental source: strain 1021, megaplasmid pSymb
R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, A.; Ampe, F.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A.Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A.Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A.Reference number: A96039; MUID:21368234; PMID:11474104
A.Contents: annotation
C.Genetics:
A.Gene: glgB2; SMb21447
A.Genome: plasmid
C.Superfamily: trehalose trehalohydrolase
C.Keywords: glycosyltransferase; hexosyltransferase

Query Match 26.8%; Score 792.5; DB 2; Length 601;
Best Local Similarity 36.9%; Pred. No. 1.9e-46;
Matches 192; Conservative 79; Mismatches 199; Indels 51; Gaps 15;
QY 3 SGGN-IEKNKGIFKLWAPYVNSVKLKSLLIPMEKNDGEGFEVEIDIEENLTYSYII 61
DB 13 SWGANFIDSDTCRFLWAPDERVDVLGGAVHKQSLDGGWFEITL-AAKAGERVCYFRL 71
QY 62 EKREIPDPASRYQLGVHDKSLI---RTDQILDLGVKVEDLIYELHVGTFSGEQN 118
DB 72 ADGTEVADPASSAQEREASGTSIVVDHAAYENQTSWRGRPEEAVISELHVGCFTPECT 131
QY 119 FKGVEKLDYLDLGLTGITELMPVAQPCGNRDGVDGVFLYAVQNTYGGPWELAKLVNRA 178
DB 132 FRAAIERPLHLAGAGITAEIMPAQFPQGVGMGYDGLVLYAPHNAYKPDOLKALVDAA 191
QY 179 HRRGTAVIDVYVNHGPGNLYLGLGP-YFSDRYKTPWGLTFNFDRCDDQVRKFIEN 237
DB 192 HSLGTLVLDVYVNHGPGNLYSRYASRFFNDRPTMGASTAFEE---EAVRYTFIEN 248
QY 238 VEYWFRTKIDGLRDVAHAFDNPSPKHL---QBIAEK-AHQLGKGFVIAESDLNDPKI 292
DB 249 ALYWLGHFRFDGLRDATQIRDTTKPHELVALEHEVREAFARQIHLVLEDAHRRSLL 308
QY 293 VKDDCGYKI--DAQWDDPHAVHAFITKEKYVYODF-----GRTEIEKTFKDVVVD 345
DB 309 QRDASGARMFLDAAWMDLHNLHVATGETKGHYLFADEPWG---IRSALEAFVAP 365
QY 346 GYSRYRGTHGAPVD---LPRKFVVFIONHDQVGNRNGERLSILTDKTYLMAATL 402
DB 366 AKEDNF-----SPEGSRAVRVPPQGRVNFQNHQDQIGNAFGERLASLQEDSLRLAAM 419
QY 403 YILSPYIPLIFMGEYETNPFFFSDFDPVLKIGVREGRL-----KENNOM 450
DB 420 HMLTQIPLIFMGEYETNPFFFSDFDYOGEI-AAAIRLGRDEAENFGCLPEGKTMDL 478
QY 451 IDPQSEAEFLKSLKSWK-----IDEVLDYKQLINIRKRY 486
DB 479 PDPPLDDPVFAGSKLRNRRATSPAGERHLAYVRDLAVIRQRH 519
RESULT 14
S19134
probable alpha-glucanotransferase - Anabaena variabilis

C.Species: Anabaena variabilis
C.Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 17-Mar-2000
C.Accession: S19134
R:Sato, N.
Plant Mol. Biol. 18, 165-170, 1992
A.Title: Cloning of a low-temperature-induced gene lti2 from the cyanobacterium Anaba
A.Reference number: S19133; MUID:92119230
A.Accession: S19134
A.Status: preliminary
A.Molecule type: DNA
A.Residues: 1-552 <SAT>
A.Cross-references: EMBL:X59279; NID:g39252; PIDN:CAA41970.1; PID:g39253
C.Superfamily: neopullulanase; alpha-amylase core homology

Query Match 16.7%; Score 493; DB 2; Length 552;
Best Local Similarity 28.6%; Pred. No. 5.3e-26;
Matches 159; Conservative 76; Mismatches 142; Indels 178; Gaps 26;
QY 15 FKLWAPYVNSVKLKS---KKLIPMEKNDGEGFE--VEIDIEENL-----TYSYIIEDK 64
DB 7 FTLFAPYKNGAALIASFSQWQEIIPMKKGDDGYFRTTVELEDCTGYQKFRVQTRSFEEED 66
QY 65 R-EIPDPASR--YQPLGVHDKSQLRTDYQILDGLGKVKIED-----LIYELHVGT 112
DB 67 QWVDVTPYATDIDESSGKDNSTARIKDGKIVDTYVWQHDDKPLPADHELVIYELHVG 126
QY 113 FS-----QEGNFKGVIKLDYLDLGLTGITELMPVAQPCGNRDGVDGVFLYAVQNTY 165
DB 127 FSGGDDPYARGKYKHVIEKLDYLCLEGINAIELLPVKRPGDYSGWYNPRYFFATESY 186
QY 166 GGPWELAKLVNNAHKGRTAGIADLVVYVNH---SDLNDRPKIVKDCGKIDAOVDDFHHA 205
DB 187 GSTADLKLVDCHQRGIRLINDGIYNHSEASSPLTQIDHDYWHHEPRDPNN----- 240
QY 206 PYFSORYKTPWGLTFNFD--DRGCD--QVRKFEILENVEWFTFKIDGLRLDAVHAFDN 261
DB 241 -----WGPEFNHYEDENLETYPARKFIDGTVRYWVGVEYHLDGRYDAAARQIANY 290
QY 262 SPKH-ILQEIATKAHQLGKVFIAE-----SDLNDRPKIVKDCGKIDAOVDDFHHA 312
DB 291 DFMHWTQAQAKTAGAKPNVNAEHIPETTSITNLDGP-----MDGCWHHSFYHT 340
QY 313 VHAFITKEDYVYQDFGRIDIEKTFKDVVYDGKYSRYRGTHGAPVGLDLPKRFVFEI 372
DB 341 IRAHIC-----GDTFDELE-NLKQVI--DPRKQGLGATN-----VVNVL 376
QY 373 QNHD-----QVGNRNGERLSILTDK-----TTLMAATLYILSPYIPLIFMGEY 418
DB 377 TNHDDHIMVELGNR-----EIFHDEAFPRAKLGTAILMTAV-----GVPLIMWGEF 424
QY 419 YETNPFFFSDFSDPVLKIGVREGRLKENNOMIDPQSEAEFLKSLW-----KIDEEVL 473
DB 425 GEYKP-----KQDDQ-----SKIDTWLLGNDLRSFL 451
QY 474 DYVYKQLINIRKRYNN 488
DB 452 DYHKGLIGLRK--NN 464

RESULT 15
AH1915
hypothetical protein all0875 [imported] - Anabaena sp. (strain PCC 7120)
C.Species: Anabaena sp.
A.Note: Anabaena sp. (strain PCC 7120) is a synonym of Nostoc sp. strain PCC 7120
C.Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 11-Jan-2002
C.Accession: AH1915
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Irigu
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata
DNA Res. 8, 205-213, 2001
A.Title: Complete genomic sequence of the filamentous nitrogen-fixing Cyanobacterium
A.Reference number: AB1807; MUID:21595285; PMID:11759840
A.Accession: AH1915

A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-552 <KUR>
A:Cross-references: GB:BA000019; PIDN:BA072832.1; PID:g17130220; GSPDB:GN00179
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: all0875
C:Superfamily: neopullulanase; alpha-amylase core homology

Query Match 16.78; Score 493; DB 2; Length 552;
Best Local Similarity 28.68; Pred. No. 5.3e-26;
Matches 159; Conservative 76; Mismatches 142; Indels 178; Gaps 26;

QY 15 FKLWAPVNSVKLKL---KKLIPMEKNDGFFE--VEIDDIEENL--/-TYSYIIEDEK 64
DB 7 FTLFAPYKGAALIASDSQOEIPMKKGGDGYFTTVELEDGYQYKFRVQTRSWFFED 66
QY 65 R--EIPDPASR--YQPLGVHDKSQLIRTYQILDGLGVKIED-----LIYELHVGT 112
DB 67 QWVDVTPYATDIDESSGKDSIARIKDGKIVDYYVWQHDDKPLPADHELVIYELHVGD 126
QY 113 FS-----QEGNFKGVIEKLDYKLDGICIELMPVAQPGNDRDGYDGVFLYAVONTY 165
DB 127 FSGEDDPYARGYKHYKHYKHYKHYKHYKHYKHYKHYKHYKHYKHYKHYKHYKHYKHY 186
QY 166 GGPWEAKLVNEAKHKGIAVILDVVYVYH-----GPEGNYLLGLG 205
DB 187 GSTADLKLKLVDECHQIRIIMDGIYVHSEASSPLTQIDHDYVWHHEPRDPDN----- 240
QY 206 PYFSDRYKTPWGLTFNFD--DRGCD--QVRKFILENVEYWFKFKIDGLRLDAVHAIFDN 261
DB 241 -----WGPEFNYEHVDENLETYPARKFIGDTVRYWVGEVHLDGIRYDAARQIANY 290
QY 262 SPKH-ILQETAERKAHQKGFVIAB-----SDLNDPKIVKDDCGYKIDAQWVDDPHHA 312
DB 291 DFHWIAQEAQKTAGAKPFYNVAEHIPEITTSITNLDGP-----MDGCWHDSEYHT 340
QY 313 VHAFITREKDYVYQDFGRIDIEKTFKDVFDGKYSRYGRTHGAPVGDLPKRKFVVEI 372
DB 341 IKAHIC-----GDTFDE-ILKADVI--DPKROGFLGATN-----VVNYL 376
QY 373 QNH-----QVGNRNGERLSILTDK-----TTLMAATLYILSPYIPIFMGEY 418
DB 377 TNHDDHIMVELGNR-----EIFHDEAFRAKLGTAIIMTAV-----GVPLIWMGEF 424
QY 419 YETNPFFSDFSDPVLKGVREGRLENQMDPQSEEAFLKSLW-----KIDEEVL 473
DB 425 GEYKP-----KQDQ-----SKIDWTLGNDLNRSLF 451
QY 474 DYKQOLINIRKRYNN 488
DB 452 DYHKGILGLRK--NN 464

Search completed: July 15, 2002, 12:18:44
Job time: 136 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 12:24:32 ; Search time 19.24 Seconds
(without alignments)
1118.923 Million cell updates/sec

Title: US-09-298-924-8
Perfect score: 2954
Sequence: 1 MFSPGGNIKNKGIFKLWAP.....KLKKDELIKVNRGVGVYQLQ 556

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	969	32.8	596	1 TREZ_RHISP	Q53238 rhizobium s
2	959	32.5	598	1 TREZ_ARTSP	Q44316 arthrobacte
3	936	31.7	575	1 TREZ_ARTRM	Q94j66 arthrobacte
4	890.5	30.1	589	1 TREZ_BREHE	O52520 brevbacte
5	866.5	29.3	580	1 TREZ_MVCTU	Q10769 mycobacteri
6	353	11.9	666	1 GLGB_BACCL	P30537 bacillus ca
7	342.5	11.6	639	1 GLGB_BACST	P30538 bacillus st
8	320	10.8	843	1 PUA_A_THEMA	O33840 thermotoga
9	315	10.7	562	1 AMY2_DICTH	P14898 dictyoglomu
10	313.5	10.6	627	1 GLGB_BACSU	P39118 bacillus su
11	296.5	10.0	764	1 GLGB_STRAU	P52980 streptomyce
12	282	9.5	770	1 GLGB_SYNY3	P52981 synecocyst
13	276.5	9.4	659	1 GLGX_HAEIN	P45178 haemophilus
14	273	9.2	731	1 GLGB_MVCTU	Q10625 mycobacteri
15	270	9.1	702	1 GLGB_HUMAN	Q04446 homo sapien
16	269	9.1	734	1 GLGB_AGRTU	P52979 agrobacteri
17	267	9.0	639	1 GLGB_BUTFI	P30539 butyrivibri
18	266	9.0	730	1 GLGB_HAEIN	P45177 haemophilus
19	265	9.0	520	1 AMY_BACME	P20845 bacillus me
20	264.5	9.0	773	1 GLGB_SYNP7	P16954 synecococc
21	262.5	8.9	498	1 AMY3_DICTH	P14899 dictyoglomu
22	261.5	8.9	558	1 O16G_BACCE	P21332 bacillus ce
23	257	8.7	561	1 TRFC_BACSU	P39795 bacillus su
24	254.5	8.6	562	1 O16G_BACTR	P29094 bacillus th
25	252.5	8.5	512	1 AMY1_DEBOC	P19269 debaryomyce
26	250.5	8.5	1196	1 AMYB_FAEPO	P21543 paenibacill
27	249	8.4	588	1 NEPU_BACST	P38940 bacillus st
28	243.5	8.2	478	1 YDD1_SCHPO	Q10427 schizosacch
29	243	8.2	561	1 O16G_BACSU	Q06994 bacillus su
30	239.5	8.1	574	1 CDAS_THEET	P29964 thermoanaer
31	237	8.0	494	1 AMY1_SACFI	P21567 saccharomyc
32	234	7.9	555	1 O16G_BACCO	Q45101 bacillus co
33	234	7.9	1475	1 APU_THEY	P16950 t amylopull

34	231	7.8	1481	1 APU_THEET	P38939 t amylopull
35	229.5	7.8	508	1 O16G_BACSP	P29093 bacillus sp
36	228	7.7	799	1 GLGB_MAIZE	Q08047 zea mays (m
37	225.5	7.6	498	1 AMYA_ASPAW	Q02805 aspergillus
38	225.5	7.6	499	1 AMYA_ASPOR	P10529 aspergillus
39	225.5	7.6	499	1 AMYB_ASPAW	Q02906 aspergillus
40	225.5	7.6	569	1 MALT_CANAL	Q02751 candida alb
41	224.5	7.6	535	1 DEXB_STRPN	Q54796 streptococc
42	224	7.6	586	1 AMYM_BACAD	P32818 bacillus ac
43	224	7.6	728	1 GLGB_ECOLI	P07762 escherichia
44	223.5	7.6	499	1 AMY_ASPSH	P30292 aspergillus
45	223.5	7.6	777	1 ISOA_FLASP	O32611 flavobacter

ALIGNMENTS

RESULT	1
TREZ_RHISP	
ID	TREZ_RHISP
AC	Q53238
DT	16-OCT-2001 (Rel. 40, Created)
DT	16-OCT-2001 (Rel. 40, Last sequence update)
DT	16-OCT-2001 (Rel. 40, Last annotation update)
DE	Malto-oligosyltrehalose trehalohydrolase (EC 3.2.1.141) (MTHase) (4-
DE	alpha-D-[(1->4)-alpha-D-glucano]trehalose trehalohydrolase)
DE	(Maltooligosyl trehalose trehalohydrolase).
GN	TREZ.
OS	Rhizobium sp. (strain M-11).
OC	Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC	Rhizobiaceae; Rhizobium.
OX	NCBI_TaxID=391;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=96219094; PubMed=8829547;
RA	Maruta K., Hattori K., Nakada T., Kubota M., Sugimoto T., Kurimoto M.;
RT	"Cloning and sequencing of trehalose biosynthesis genes from Rhizobium
RT	sp. M-11."
RL	Biosci. Biotechnol. Biochem. 60:717-720(1996).
CC	- - CATALYTIC ACTIVITY: Hydrolysis of alpha-(1->4)-D-glucosidic
CC	linkage in 4-alpha-D-[(1->4)-alpha-D-glucanosyl](n) trehalose to
CC	yield trehalose and alpha-(1->4)-D-glucan.
CC	- - PATHWAY: Trehalose biosynthesis.
CC	- - SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC	KNOWN AS THE ALPHA-AMYLASE FAMILY
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC	or send an email to license@isb-sib.ch).
CC	-----
CC	EMBL; D78001; BAA11187.1; -
DR	InterPro; IPR000461; Alpha-amyase.
DR	Pfam; PF00128; alpha-amyase; 1.
DR	Hydrolase; Glycosidase.
FT	ACT_SITE 265 265 BY SIMILARITY.
SQ	SEQUENCE 596 AA; 65262 MW; 73E8A0AE0534DDCD CRC64;

Query Match	32.8%;	Score	969;	DB 1;	Length	596;	
Best Local Similarity	39.5%;	Pred. No.	1e-60;	Indels	60;	Gaps	12;
Matches	215;	Conservative	74;	Mismatches	195;		
QY	13	GFIFKLWAPVNSVKLSKLIPMEKN-----DEGFFEVEIDDIENLTYSYIIE-DKR	65				
Db	15	GRFDIWAPEGTVTLLAGGERVEMGRPCNGPADEGWTAADAPTADVDYGYLLDGDDEI	74				
QY	66	ETPDASRYOPVGLGVHDKSOLIRTDYQILDGLCKVKIEDL-----IIVELHVGTFSQ	115				
Db	75	PLPDPRTRRQPEGVH---ALSRT-----FDPGAHRWQDAGWQRELQGSVIYELHIGTFP	127				

```

QY 116 EGNEFGVIEKLDYKDLGITGIELMPVAQFPGNDRMGVDFVFLYAVONTYGGPWELAKLV 175
D 128 EGTDAAGAKLDYLAGLIDFIELLPVNAFNGTHWGYDGVQWFAVHEGYGPAAYORFV 187
QY 176 NEAHKRGTAVIDVYNNHIGPEGNVLGLGYPFSDRYKTPWGLTFNFDRCDOVRKFL 235
D 188 DAHAAGLVYDQVYNNHIGPEGNVLGYPFSDRYKTPWGLTFNFDRCDOVRKFL 247
QY 236 ENVEYWFRTKIDGLRLDAVHAIFDNPSPKHILQIETAEKAKHOLGK-----FVIAESDLN 289
D 248 DNVAWLRDVRVGLRLDAVHALKDERAVHILIEFGALADALSSEGGRLPLTIAESDLN 307
QY 290 PKIV--KDDCYKIDAQVDDFHHAVHAFITKEKDYDQFGRIEDIEKTFKQVVFYDGG 347
D 308 PRLLYPRDNGYGLAGWSDDFHHAVHVNVSGETTGYYSDFSLGALAKLVLRDGFHDS 367
QY 348 YSRVGRTHGAPV--GDLPPRKFFVFTONHDQVNRNGERLSLTLDKTYLMAATLYIL 405
D 368 YSSFRGCHGRPINFSVHPAALVVCSONHDQIGNRATGDRLSQSLPYGSLALAAVLTIT 427
QY 406 SPYIPLFMGEYETNPFFSDFSDPVLTKGVREGRLKNNOM-----IDPQSEEA 458
D 428 GPFTPLFMGEYCATTPWQFFTSHPPELKGATAEGRIREFRGMGNDPAVVPDPQDPET 487
QY 459 FLKSKLSWK-----IDEVLDYKOLINIRKRYNCKRVKVRREGNCITLIMEKIGIIA 513
D 488 FTRSKLDWAESAGDHARLLELYRSLTLR-----RSTPELRLGFDATAV----- 533
QY 514 SFDD 517
D 534 EFDD 537

RESULT 2
TREZ_ARTSP
ID TREZ_ARTSP STANDARD; PRT; 598 AA.
AC Q44316;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Malto-oligosyltrehalose trehalohydrolase (EC 3.2.1.141) (MTHase) (4-
DE alpha-D-[(1->4)-alpha-D-glucano]trehalose trehalohydrolase)
DE (Malto)oligosyl trehalose trehalohydrolase).
GN TREZ.
OS Arthrobacter sp. (strain Q36).
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Micrococciaceae; Micrococciaceae; Arthrobacter.
OX NCBI_TaxID=1667;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96195835; PubMed=8605217;
RA Maruta K., Hattori K., Nakada T., Kubota M., Sugimoto T., Kurimoto M.;
RT "Cloning and sequencing of trehalose biosynthesis genes from
RT Arthrobacter sp. Q36.;"
RL Biochim. Biophys. Acta 1289:10-13(1996).
CC -1- CATALYTIC ACTIVITY: Hydrolysis of alpha-(1->4)-D-glucosidic
CC linkage in 4-alpha-D-[(1->4)-alpha-D-glucano]trehalose to
CC yield trehalose and alpha-(1->4)-D-glucan.
CC -1- PATHWAY: Trehalose biosynthesis.
CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: D63343; BAA09668.1; -
DR InterPro; IPR00461; Alpha_amylase.

```

```

DR InterPro; IPR004193; isoamylase_N.
DR Pfam; PF00128; alpha-amylase; 1.
DR Pfam; PF02922; isoamylase_N; 1.
KW Hydrolase; Glycosidase.
FT ACT_SITE 267 BY SIMILARITY.
SQ SEQUENCE 598 AA; 65831 MW; 8B5C610AD3766947 CRC64;

```

```

Query Match 32.5%; Score 959; DB 1; Length 598;
Best Local Similarity 40.2%; Pred. No. 5.1e-60;
Matches 202; Conservative 81; Mismatches 186; Indels 34; Gaps 10;

```

```

QY 15 FKLHAPVYNSVKLSKSLIPMEK-----NDEGFEVEIDDEENITYKILP-DKREI 67
D 19 YDWAPNAESVTLTLAGGERYAMQRAETPEDAGWTAAGAPTDCNVDYGLDGETPL 78
QY 68 PDASRYQPLGVHDKSOLI--RTDQIILD---LGRVKIEDLIIIVELHVGTSQEGNFGVI 123
D 79 PDPRTRQPGVGHALSRTFDPDSAYSQDDAWQGR-ELQGAIVYIELHGTFTPEGTLEAAA 137
QY 124 EKDLYLKDLGITGIELMPVAQFPGNDRMGVDFVFLYAVONTYGGPWELAKLVNEAHRGI 183
D 138 GKLDYLAGLVDFTELLPVNAFNGTHWGYDGVQWFAVHEAGGPEAYQRFVDAHAAGL 197
QY 184 AVLDVYNNHIGPEGNVLGLGYPFSDRYKTPWGLTFNFDRCDOVRKFLILENVEYWF 243
D 198 GVTDVYNNHIGPEGNVLGYPFSDRYKTPWGLTFNFDRCDOVRKFLILENVEYWF 257
QY 244 TFKIDGLRLDAVHAIFDNPSPKHILQIETAEKAKHOLGKVF-----IAESDLNDPKIV--KD 295
D 258 DYRVGDLRLDAVHALKDERAVHILIEFGALADQISAEVGRPLTLIAESDLNPKLLYPRD 317
QY 296 DCGYKIDAQVDDFHHAVHAFITKEKDYDQFGRIEDIEKTFKQVVFYDGGYSRYGRGT 355
D 318 VNGYGLGWSDDFHHAVHVNVTGETTGYYSDFDLSAALAKLVLRDGFHDSYSSFRERH 377
QY 356 HGAPV--GDLPPRKFFVFTONHDQVNRNGERLSLTLDKTYLMAATLYILSPYPLIF 413
D 378 HGRPINFSVHPAALVVCSONHDQIGNRATGDRLSQSLPYGSLALAAVLTGFTPMLL 437
QY 414 MGEYETNPFFSDFSDPVLTKGVREGRLKNNOM-----IDPQSEAFKLSLW 466
D 438 MGEYETNPFFSDFSDPVLTKGVREGRLKNNOM-----IDPQSEAFKLSLW 497
QY 467 KIDEE-----VLDYKOLINIRK 484
D 498 AEAEGDHARLLELYRSLTLR 520

```

```

RESULT 3
TREZ_ARTSM
ID TREZ_ARTSM STANDARD; PRT; 575 AA.
AC Q9AJN6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Malto-oligosyltrehalose trehalohydrolase (EC 3.2.1.141) (MTHase) (4-
DE alpha-D-[(1->4)-alpha-D-glucano]trehalose trehalohydrolase)
DE (Malto)oligosyl trehalose trehalohydrolase).
GN TREZ.
OS Arthrobacter ramosus.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Micrococciaceae; Micrococciaceae; Arthrobacter.
OX NCBI_TaxID=1672;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S34;
RA Yamanoto T., Maruta K., Watanabe H., Yamashita H., Kubota M.,
RA Fukuda S., Kurimoto M.;
RT "Trehalose producing operon trez from Arthrobacter ramosus S34."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: Hydrolysis of alpha-(1->4)-D-glucosidic
CC linkage in 4-alpha-D-[(1->4)-alpha-D-glucano]trehalose to

```



```
Db 414 ILPSHDEVHV-GKSLLSKMPGTYEKFQRLRLLYGVLTHPGCKLLFMGGERGQ----- 468
Qy 425 FFFSDFSPVLKGVREGRLKENNOMIDPQSEAEFLSKLSWKIDEEVLDYKOLINRK 484
Db 469 -FDEWKD-----LQOLDWML-----FDFDMHRNNMNVKELLCKYK 503
Qy 485 RY 486
Db 504 RY 505

RESULT 8
PULA_THEME
ID PULA_THEME STANDARD; PRT; 843 AA.
AC O33840;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Pullulanase precursor (EC 3.2.1.41) (Alpha-dextrin endo-1,6-alpha-
DE glucosidase) (Pullulan 6-glucanohydrolase).
GN PULA OR TM1845.
OS Thermotoga maritima.
OC Bacteria; Thermotogales; Thermotoga.
OX NCBI_TaxID=2336;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MSB8 / DSM 3109;
RX MEDLINE=98115241; PubMed=9453151;
RA Bibel M., Brett C., Gossler U., Kriegshauser G., Liebl W.;
RT "Isolation and analysis of genes for amylolytic enzymes of the
RT hyperthermophilic bacterium Thermotoga maritima.";
RL FEMS Microbiol. Lett. 158:9-15(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MSB8 / DSM 3109;
RX MEDLINE=99287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA Hart D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT "Evidence for lateral gene transfer between Archaea and Bacteria from
RT genome sequence of Thermotoga maritima.";
RL Nature 399:323-329(1999).
CC -1- CATALYTIC ACTIVITY: Hydrolysis of (1->6)-alpha-D-glucosidic
CC linkages in pullulan and in amylopectin and glycogen, and the
CC alpha- and beta-limit dextrins of amylopectin and glycogen.
CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ001087; CAA04522.1; -
DR EMBL; AE001821; AAD36907.1; -
DR TIGR; TM1845; -
DR InterPro; IPR000461; Alpha_amylase.
DR InterPro; IPR004193; Isoamylase.N.
DR Pfam; PF00128; alpha-amylase; 1.
DR Pfam; PF02922; isoamylase.N; 1.
KW Hydrolase; Glycosidase; Signal; Complete proteome.
FT SIGNAL 1
FT CHAIN 19 POTENTIAL.
FT ACT_SITE 20 843 PULLULANASE.
FT ACT_SITE 535 535 BY SIMILARITY.
FT ACT_SITE 564 564 BY SIMILARITY.
FT ACT_SITE 652 652 BY SIMILARITY.
SQ SEQUENCE 843 AA; 96261 MW; C42DDE233D54FE77 CRC64;
```

```
Query Match 10.8%; Score 320; DB 1; Length 843;
Best Local Similarity 23.4%; Pred. No. 6.3e-15;
Matches 143; Conservative 88; Mismatches 185; Indels 194; Gaps 31;

Qy 5 GGNIEKNKGIFKLMAPYVNVSKLKS-----KLPMEKNDGEFFVEIDDIENLT 56
Db 226 GAVYSPEKTIFRVWSPYKWKVLLFKNGEDTEPYQVYVNMVEYKNGVWEAVGDLGVF 285
Qy 57 YSIIEDK---REIPDPASRQPLGVHDKSQ-----LIRTDQ--LIDLQ-KVK-TEDL 103
Db 286 YLQLENYGKIRTVDPYSK-----AVYANKSKSAVWLARTNPGCWENDRGPKIEGEDA 341
Qy 104 IYELHVGTF-----QEGNFK---GVIEKLDYKLDIGITGIELMPVAQ 144
Db 342 IYELHITADITGLNSGVKNKGLYLGLTEENTKPGGVTTGLSHLVGLVTHVHILPFED 401
Qy 145 F-PGNR-----DWGYDGVFLYAV-----ONTYGGPWELAKLYNEAHKRGIAV 185
Db 402 FYTGDELDKDFEKYYNNGYD-PYLFMVPEGRYSTDPKNPHTRIREVKEVMVKALHKGIGV 460
Qy 186 ILDVVYVNHIGPEGNYLLGLG-----PYESDRY-KTPWGLTFNFDRCGQOV----- 230
Db 461 IMDVFPFH-----TYGIGELSAFDQTVPIYFYRIDKT--GAYLN--ESGCGNVIASER 509
Qy 231 ---RKFILENVYWFKTFKIDGLRLDAVHAIFONSPRKHILQIEAIAEKALQKGFVIAESDL 287
Db 510 PMRKRFIVDTVYVWKEYHYDGFREFDOMGLI---DKTMELEVERALHKI----- 555
Qy 288 NDPKIV-----KDCGKYKIDAQWVDFHHAHV--AFITKEKDYVYODFG 329
Db 556 -DPTIILYGEPPGWGAGPIRFGKSDVAGTHVAFAFNDEFDAIRGVSFNPVSKGFVMSGYG 614
Qy 330 RIEDIEKTFKDFVYDGKYSRYGRTHGAPVGDLPFRKFVFEIQNHQDVGNRNGERLSI 389
Db 615 KETKIKRGVGSINYDGKLIKSPA-----LDPETINYAACHDN-----HT 655
Qy 390 LTKDTTYL-----MAATLYILSPYIPLIFMGREYETNPFFPFS 429
Db 656 LWDK-NYLAAKADKKKEWTEELKNAQKLAGAILTSQGVFPLHGGGDFCRKTN-FDNS 713
Qy 430 FSDPVLKGVREGRLKENNOMIDPQSEAEFLSKLSWKIDEEVLDYKOLINRK----- 484
Db 714 YNAPISNG-----FDYERKLQFI-----DVFNYHKGLIKRKEHPAF 751
Qy 485 RYNNKRVKE 494
Db 752 RLKNAEIEKK 761

RESULT 9
AMY2_DICTH
ID AMY2_DICTH STANDARD; PRT; 562 AA.
AC P14898;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-DEC-1992 (Rel. 24, Last annotation update)
DE Alpha-amylase 2 (EC 3.2.1.1) (1,4-alpha-D-glucan glucanohydrolase).
GN AMYB.
OS Dictyoglomus thermophilum.
OC Bacteria; Dictyoglomus group; Dictyoglomus.
OX NCBI_TaxID=14;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-8.
RC STRAIN=H-6-12;
RX MEDLINE=88329076; PubMed=2458257;
RA Horinouchi S., Fukusumi S., Ohshima T., Beppu T.;
RT "Cloning and expression in Escherichia coli of two additional amylase
RT genes of a strictly anaerobic thermophile, Dictyoglomus thermophilum,
RT and their nucleotide sequences with extremely low
RT guanine-plus-cytosine contents.";
RL Eur. J. Biochem. 176:243-253(1988).
```



```
Db 26 FGSHTRELNGSGYEFVWAPASVRVAGDPSNGSGEHHVRHVRNDNGINTFLFPGIGE 85
Qy 54 NLUTSY-IIEKREI---PDPAASYOPLGVH-----DKSQLIRTDYQ 91
Db 86 KERYKYEIVTNGEIRLAKDPYAIYSEVRPNTASLTLDLEGYSWQDKWQKQAKTIYE 145
Qy 92 ILDLGVKIEDLIYELHVGTFSGRG-----NFKGVIEKL-DYLDKDLGITGIELMPVAQF 145
Db 146 -----KPVFIYELHLSGSKHSDGRHSYKLSQTLPIYIKKHGFTHEILLPVYEH 196
Qy 146 PGNRDWGVGYFLYAVQNTYGGPWEAKLVNEAKHKGITAVILDVYVNIH--GPESNYLLG 203
Db 197 PYDRSGIOGTGYSPTRSGFPPHDLMKFVDECHOQNGVILDWVPGHCFKDAHGLYMF 256
Qy 204 LGPYFS-----DRYKTPWGLTFNFDRCDCOVKRFILNVEYWFKTFKIDGLRLDAVHAI 258
Db 257 GEPLYEYKEERDENLWLG-TANF-DLGKPEVHSLISNLYWAEFYHIDGFRVDVANI 314
Qy 259 F---DNSPKH---LQETAERAKHQLGK-----FVIAESDLNDPKI---VKDDCG-----YK 300
Db 315 LYWPNQDERHNPYAVDPLKLNQTMREAYPHVMMAEDSTEWQVTVGAVEEGGLGPHYK 374
Qy 301 IDAOWVDDFHAVHAFITKEDYQ--DFGRIEDIEKTFKDVFYD-----GKYSRYRGR 354
Db 375 WNMGMNDVLXYMET-PPERRHCHQLISFSLLYAFSEHFVLPFSDHDEVYVYKKSLL-----429
Qy 355 THGAPVGLDPKRFVVFQNHQDGVNGRNGERLSILTDTKTYLMAATLYILSPYIPLIFM 414
Db 430 -----LNKMPGDYQKQFAQ-----YRLLGLVMTVHPGKLLIFM 462
Qy 415 GEYYETNPFFSDFSDPVLKIGVREGRLKNNQMD--POSBEA-----FLKS 462
Db 463 GSEFAQ-----FDEWKD-----TEQLDWFLDSPFMHOKASVFTQDLRLRFYOKS 505
Qy 463 KLSWKIDE-----EVLDDY--KOLINIRKRYNNCKRVEKREGNCITLINEKIGITASF 515
Db 506 KILYEHDIRAOSFEWIDVHNDQSFISFIRY-----QKHG-----EALVVICNF 550
Qy 516 DDIVNSKITG 526
Db 551 TPVVVHYQVDVG 561

RESULT 11
GLGB_STRAU STANDARD; PRT; 764 AA.
AC P52980;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE 01-OCT-1996 (Rel. 34, Last annotation update)
DE 1.4-alpha-glucan branching enzyme (EC 2.4.1.18) (Glycogen branching
enzyme).
GN GLGB.
OS Streptomyces aureofaciens.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1894;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CCM 3239 / ATCC 10762;
RA MEDLINE=94347823; PubMed=8068720;
RT "Cloning of the putative glycogen branching enzyme gene, glgB, from
RL Biochim. Biophys. Acta 1200:334-336(1994).
CC -1- CATALYTIC ACTIVITY: Formation of 1,6-glucosidic linkages of
CC glycogen.
CC -1- PATHWAY: THIRD STEP IN GLYCOGEN BIOSYNTHESIS.
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC -----
```

```
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed, usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L11647; AAA67437.1; -.
DR InterPro; IPR000461; Alpha_amylase.
DR InterPro; IPR001899; Gram_pos_anchor.
DR InterPro; IPR004193; Isoamylase_N.
DR Pfam; PF00128; alpha-amylase; 1.
DR Pfam; PF02922; isoamylase_N; 1.
KW Glycogen biosynthesis; Transferase; Glycosyltransferase.
FT ACT_SITE 440 440 BY SIMILARITY.
FT ACT_SITE 493 493 BY SIMILARITY.
FT ACT_SITE 561 561 BY SIMILARITY.
SQ SEQUENCE 764 AA; 85325 MW; 6B45482B4A368ACF CRC64;

Query Match 10.0%; Score 296.5; DB 1; Length 764;
Best Local Similarity 24.6%; Pred. No. 2:5e-13;
Matches 125; Conservative 88; Mismatches 177; Indels 119; Gaps 27;

Qy 3 SFGNGIEKNKGI-----FKLWAPYVNSVKLKSKKL-----IPMEK-NDEGFFEVEIDD 50
Db 162 ALGSOPMEHQGVAGTRFTVWAP--NALGVRTGDFSYWDVAVAYPMRSLGASGVWELFLPG 219
Qy 51 IEENLTSYII---EDKREI-PDPASRYOPLGVHDKLSQLIRTDYQIIDL-----GKVKI 100
Db 220 VAEGALYKYEITRDPDGGTTLRADPMARYAEVPPANASIVTASRYEWQDAENMARRALAP 279
Qy 101 ED--LIIVELHVGTFSGEGNFKGVIEKLD-YLKDLGITGIELMPVAOPPGNRDNGYDGVF 157
Db 280 HQAPMSVVELHLSWRGCLSYRLAEQLPAYKELGTHVLMELPVAEHPGSGSGYQVTG 339
Qy 158 LYAVONTYGGPWEAKLVNEAKHKGIAVILDVYVNIHGPENYLLGLGPYFSDR-----YK 213
Db 340 FYAPTSMGTPTDDPRFLVDALHRAIGIVIVDWVPAHF-PRDDWALA---EFDGRPLYEHQ 395
Qy 214 TP-----WCLTFNFDRCDCOVKRFILNVEYWFKTFKIDGLRLDAVHAFI-----259
Db 396 DPRAAHPDWG-TLEF-DYGRKVRNFLVANAVYWCQEFHVDGLRADAVASMLYLDYSRD 453
Qy 260 --DNSPK-----HILQETAERAKHQL--GKFVIAESDLNDPKIVK--DDCGYKID 302
Db 454 EGDWSPNAHGGREDLDAVALLQENNAVYRFPQVVTIAEESTAWDGVTRPTDSGGLGFG 513
Qy 303 AQWVDDFHAVHAFITKE---KDYVQDFGRIEDIEKTFKDVFYVDGKYSRYGRTHGAP 359
Db 514 LKWNMGWMDTLRYVSKPEVHRKYHHDM-----TFGMVYAFS-----551
Qy 360 VGDLPPRRKVVVFQNHQDGVNGRNGERLSILT-----DKTYLMAATLYILSPYIPL 411
Db 552 -----ENFVLPI-SHDE---VHGKR-SLVSKMPCGDWQOORATHRAYLGFMAHPKQL 600
Qy 412 IFMGEEYETN-----PFFFSDFSDP 433
Db 601 LFMQGEFAQGSWSETYGPDWVWLDSSYP 629

RESULT 12
GLGB_SYNY3
ID GLGB_SYNY3 STANDARD; PRT; 770 AA.
AC P52981;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE 1.4-alpha-glucan branching enzyme (EC 2.4.1.18) (Glycogen branching
enzyme).
GN GLGB OR SLL0158.
OS Synechocystis sp. (strain PCC 6803).
```


Db 375 HGYGVGNPPSYFAEMNDRFDDLCRFLWKSGETGAFAPAGSSDLFKKNDRLPHTTL 434
Qy 337 -----TFKDVYVQGYKSYRGTRHGAPVGDLPKRFVVFQNDHQVGNRGNGRL 387
Db 435 NFITAHDGFTLKDLVSYNQKHNETNGENRGRNE-----NYSNH---GVGSESTL 484
Qy 388 -----SILTKDTYLMATL--YILSPYIPIFMGEYYETNPFFFSDFSDPVLKGR 440
Db 485 SEPOKSAVENNRTPAQSGLLMSLLLANGTPLLADGDFGNTQ----- 526
Qy 441 EGRLEKNQMDPQSEAEFLKLSW-KIDDEVLDYYKQLINIRKRYNCKR 491
Db 527 ----YGNNNAYQDNEITWLK-----WANFNEELFELTKTIALRKOIGSLNK 570

RESULT 14
GLGB_MYCTU STANDARD; PRT; 731 AA.
AC Q10625;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable 1,4-alpha-glucan branching enzyme (EC 2.4.1.18) (Glycogen branching enzyme).
GN GLGB OR RV1326C OR MT1368 OR MTCY130.11C.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S. T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D., Gordon S. V., Eiglmeier K., Gas S., Barry C. E. III, Tekaita F., Badcock K., Basham D., Brown D., Chillingworth T., Connor R., Davies K., Devlin K., Felwell T., Gentles S., Hamlin N., Holroyd S., Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L., Oliver S., Osborne J., Quail M. A., Rajandream M. A., Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R., Sutton J. E., Taylor K., Whitehead S., Barrrell B. G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence."
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R. D., Alland D., Eisen J. A., Carpenter L., White O., Peterson J., DeBoy R., Dodson R., Gwinn M. B., Haft D., Hickey E., Kolonay J. F., Nelson W. C., Umayam L. A., Ermolaeva M. D., Salzberg S. L., Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A., Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and laboratory strains."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: Formation of 1,6-glucosidic linkages of glycogen.
CC -1- PATHWAY: THIRD STEP IN GLYCOGEN BIOSYNTHESIS.
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z73902; CAA98090.1; -
DR EMBL; AE007010; AAK45632.1; -
DR TIGR; MT1368; -

DR TubercuList; RV1326C; -
DR InterPro; IPR000461; Alpha_amylase.
DR InterPro; IPR004193; isoamylase_N.
DR Pfam; PF00128; alpha-amylase; 1.
DR Pfam; PF02922; isoamylase_N; 1.
KW Glycogen biosynthesis; Transferase; Glycosyltransferase;
KW Complete proteome.
FT ACT_SITE 411 411 BY SIMILARITY.
FT ACT_SITE 464 464 BY SIMILARITY.
FT ACT_SITE 532 532 BY SIMILARITY.
FT CONFLICT 214 214 P -> A (IN REF. 2).
FT CONFLICT 223 223 P -> Q (IN REF. 2).
SQ SEQUENCE 731 AA; 81729 MW; EE2BFEF765352617 CRC64;

Query Match 9.2%; Score 273; DB 1; Length 731;
Best Local Similarity 24.7%; Pred. No. le-11;
Matches 118; Conservative 71; Mismatches 168; Indels 120; Gaps 23;

Qy 15 FKLWAPYVNSVKL-----KLSKKLIPME-KNDEGFFEVEIDDIENLTYSY-----II 61
Db 149 FAVWAPNAKGVSLIGFNGWNGHEAPMRVLGSPGVWELFWDFPCDGLYKFRVHCADGVV 208
Qy 62 EDKREIPDPASRYQPLGVHDKSLIRTDYQILD-----LGKVKIEDLIYELHVGTF 113
Db 209 TDR---ADPPAFGTEVPPTASRVTSDDYTWGDDMMAGRALRNPVNEAMSTYEVHLGSW 265
Qy 114 SOEGNFKGVIEKL-DYLDKDLGITIELMPVAQFFGNRDWGDYVFLYAVQNTYGGPWELA 172
Db 266 RGLSYRQLARELTDYIVDQGFTHVELLPVAEHPFAGSWGQVTSYAPTFRFGTDPDFR 325
Qy 173 KLVNEAHRGIATVILDVYVNHIGPEGVLLGLGVFSDRYK-TP-----WG 217
Db 326 ALVDALHQAGICVIVDWPAPHF-PKDAWALG-----RFQCTPLYEHSDFPKRGEQLDWG 377
Qy 218 LTFNFDGRGCDQVRKFLILENVEYWFKFKIDGLRLDA-----VHAI 258
Db 378 -TYVF-DGGRPEVRNFLVANALYWLQEFHIDGLRVDAVASMVLDYSRPEGGWTNVHGG 435
Qy 259 FDN-SPRHLQIEIAEKHQL--GKVFIAE-----SDLNDPKIVKDDCGYKIDAQWVDDFH 310
Db 436 RENLEAVQFLOEMNATAHKVAPGIVTAEESTPWSGVTRPTNIG---GLGFSMKWNMGWM 492
Qy 311 HAVHAFITKE---KDYVYQDFGRIEDIEKTFKDFVVDGKYSRYRGRTHGAPVGDLPKPR 367
Db 493 HTLDYVSRDPVYRSYHHH-----EMTFSMLYAFSENY----- 525
Qy 368 FVVFQIHNQDVGNRNGERLSILTKTYYLMAATLYILSPY-----IPLIFMGEY 418
Db 526 --VLPLSHDEVYH-GKGTLMGRMPG--NNHVKAAGLRSLAYQAWHPGKQLLFMGQEF 578

RESULT 15
GLGB_HUMAN STANDARD; PRT; 702 AA.
AC Q04446;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 1,4-alpha-glucan branching enzyme (EC 2.4.1.18) (glycogen branching enzyme) (Brancher enzyme).
GN GBE1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=93216700; PubMed=8463281;
RA Thon V. J., Khalil M., Cannon J. F.;
RT "Isolation of human glycogen branching enzyme cDNAs by screening complementation in yeast."

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 12:24:08 ; Search time 55.94 Seconds
(without alignments)
1719.435 Million cell updates/sec

Title: US-09-298-924-8
Perfect score: 2954
Sequence: 1 MFSPGGNIKNKGIFKLWAP.....KLKDELIKVNRGVGVYQLE 556

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2954	100.0	556	1 Q53641	Q53641 sulfolobus
2	1784.5	60.4	561	17 P95867	P95867 sulfolobus
3	1748.5	59.2	559	1 Q55088	Q55088 sulfolobus
4	1711.5	57.9	559	1 Q9UW99	Q9UW99 sulfolobus
5	1290	43.7	359	17 Q973H4	Q973H4 sulfolobus
6	942.5	31.9	581	2 Q9ADI5	Q9ADI5 streptomyces
7	865.5	29.3	583	16 Q911V1	Q911V1 pseudomonas
8	835.5	28.3	600	16 Q9RX51	Q9RX51 deinococcus
9	792.5	26.8	601	16 Q92U63	Q92U63 rhizobium m
10	645	21.8	217	17 Q973H5	Q973H5 anabaena va
11	493	16.7	552	2 Q44528	Q44528 anabaena va
12	420	14.2	440	2 Q93Q35	Q93Q35 myxococcus
13	360	12.2	652	2 Q59242	Q59242 bacillus st
14	333.5	11.3	707	16 P73608	P73608 synecocyst
15	332.5	11.3	1280	16 Q97SQ7	Q97SQ7 streptococc
16	331	11.2	1142	2 Q93U29	Q93U29 bacillus sp

17	326.5	11.1	1287	2 Q9F930	Q9F930 streptococc
18	326	11.0	718	16 Q34587	Q34587 bacillus su
19	322.5	10.9	422	2 Q59243	Q59243 bacillus st
20	319.5	10.8	718	17 P95868	P95868 sulfolobus
21	311	10.5	718	2 Q69008	Q69008 thermus sp.
22	310.5	10.5	642	16 Q97QS8	Q97QS8 streptococc
23	306.5	10.4	666	16 Q84046	Q84046 chlamydia t
24	306	10.4	825	2 Q59319	Q59319 caldocellum
25	303.5	10.3	664	16 Q928F5	Q928F5 chlamydia p
26	302.5	10.2	1165	16 Q9XX88	Q9XX88 streptococc
27	302	10.2	783	10 Q04196	Q04196 arabidopsis
28	298.5	10.1	621	2 Q93HU3	Q93HU3 rhodothermu
29	298.5	10.1	716	17 Q973H3	Q973H3 sulfolobus
30	298	10.1	713	1 Q05152	Q05152 sulfolobus
31	296.5	10.0	818	10 Q41742	Q41742 zea mays (m
32	296	10.0	630	16 Q66936	Q66936 aquifex aeo
33	295	10.0	666	16 Q9PK26	Q9PK26 chlamydia m
34	295	10.0	789	10 Q22637	Q22637 zea mays (m
35	293.5	9.9	717	16 Q9K7U5	Q9K7U5 bacillus ha
36	292.5	9.9	660	1 Q9HHC8	Q9HHC8 thermococcu
37	292.5	9.9	741	3 Q9P5P3	Q9P5P3 neurospora
38	292	9.9	1072	16 Q9K6N1	Q9K6N1 bacillus ha
39	289	9.8	720	16 Q97FP9	Q97FP9 clostridium
40	289	9.8	741	2 Q59832	Q59832 streptomyce
41	288	9.7	741	2 Q9KY06	Q9KY06 streptomyce
42	286.5	9.7	666	16 Q9KNE8	Q9KNE8 vibrio chol
43	286.5	9.7	668	2 P71095	P71095 bacteroides
44	284	9.6	733	10 Q80403	Q80403 oryza sativ
45	284	9.6	1938	2 P70983	P70983 bacillus sp

ALIGNMENTS

RESULT 1

Q53641	PRELIMINARY;	PRT;	556 AA.
ID	Q53641; O08279; O08064;		
AC	Q53641; O08279; O08064;		
DT	01-NOV-1996 (TReMBLrel. 01, Created)		
DT	01-NOV-1996 (TReMBLrel. 01, Last sequence update)		
DT	01-DEC-2001 (TReMBLrel. 19, Last annotation update)		
DE	ALPHA-AMYLASE (FRAGMENT).		
GN	TREZ.		
OS	Sulfolobus acidocaldarius.		
OC	Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.		
OX	NCBI_TaxID=2285;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Kato M., Kettoku M., Miura Y., Komeda T., Konishi Y., Shindo K.,		
RA	Kobayashi K., Iwamatsu A.;		
RL	Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=ATCC33909;		
RX	MEDLINE=97135071; PubMed=8980629;		
RA	Maruta K., Mitsuzumi H., Nakada T., Kubota M., Chaen H., Fukuda S.,		
RA	Sugimoto T., Kurimoto M.;		
RT	"Cloning and sequencing of a cluster of genes encoding novel enzymes		
RT	of trehalose biosynthesis from thermophilic archaebacterium Sulfolobus		
RL	acidocaldarius.";		
RL	Biochim. Biophys. Acta 1291:177-181(1996).		
DR	EMBL; D64131; BAA11011.1; "		
DR	EMBL; D83245; BAA11863.1; "		
DR	InterPro; IPR000461; Alpha-amylase.		
DR	InterPro; IPR004193; isoamylase_N.		
DR	InterPro; IPR003662; sub_transporter.		
DR	Pfam; PF00128; alpha-amylase; 1.		
DR	Pfam; PF02922; isoamylase_N; 1.		
DR	PROSITE; PS00217; SUGAR_TRANSPORT_2; UNKNOWN_1.		
KW	Hydrolase.		
FT	NON_TER	556	556
SQ	SEQUENCE	556 AA;	64373 MW; 4215B456C8ED4E7 CRC64;

```
Query Match      100.0%; Score 2954; DB 1; Length 556;
Best Local Similarity 100.0%; Pred. No. 4.7e-189;
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MFSGGNIEKKNGIFKLPWAPYVNSVKLKLSPKLLPMEKNDEGFEVEIDDEENLTYSYI 60
Db 1 MFSGGNIEKKNGIFKLPWAPYVNSVKLKLSPKLLPMEKNDEGFEVEIDDEENLTYSYI 60

Qy 61 IEDKREIPDPASRYQPLGVHDKSOLIRTDYQILDGKVKIEDLIYELHVGTFSEOGNFK 120
Db 61 IEDKREIPDPASRYQPLGVHDKSOLIRTDYQILDGKVKIEDLIYELHVGTFSEOGNFK 120

Qy 121 GVIEKLDYLKDLGITGIELMPVAQPPGNRDWYDGVFLYAVQNTYGGPWELAKLVNEAHK 180
Db 121 GVIEKLDYLKDLGITGIELMPVAQPPGNRDWYDGVFLYAVQNTYGGPWELAKLVNEAHK 180

Qy 181 RGIIVILDVYVNHIGPEGNYLLGLGPFSDRYKTPWGLTFNFDRCQDVRFKILENVEY 240
Db 181 RGIIVILDVYVNHIGPEGNYLLGLGPFSDRYKTPWGLTFNFDRCQDVRFKILENVEY 240

Qy 241 WFKTFKIDGLRLDAVHAIFDPSPHILQEIATKAHQLGKFKFVIAESDLNDPKIVKDDCGYK 300
Db 241 WFKTFKIDGLRLDAVHAIFDPSPHILQEIATKAHQLGKFKFVIAESDLNDPKIVKDDCGYK 300

Qy 301 IDAQWVDDFHAVHAFITKEKDYIYQDFGRIDIEKTFKDVYVYDGYKSYRGRTHGAPV 360
Db 301 IDAQWVDDFHAVHAFITKEKDYIYQDFGRIDIEKTFKDVYVYDGYKSYRGRTHGAPV 360

Qy 361 GDLPPRFVFIQNHQDVGNRNGERLSILTDKTYLMAATLYILSPYIPLIFMGEEY 420
Db 361 GDLPPRFVFIQNHQDVGNRNGERLSILTDKTYLMAATLYILSPYIPLIFMGEEY 420

Qy 421 TNPEFFSDFSDPVLKGVREGRLKNNQMDPOSEAFKSKLSWKIDEBVLDYKQLI 480
Db 421 TNPEFFSDFSDPVLKGVREGRLKNNQMDPOSEAFKSKLSWKIDEBVLDYKQLI 480

Qy 481 NIKRYNNCKRKEVREGNCITLIMEKIGIIASFDDIVINSKITGNLLIGIGFPKKLKK 540
Db 481 NIKRYNNCKRKEVREGNCITLIMEKIGIIASFDDIVINSKITGNLLIGIGFPKKLKK 540

Qy 541 DELIKVNRGVGYOLE 556
Db 541 DELIKVNRGVGYOLE 556

RESULT 2
P95867 PRELIMINARY; PRT; 561 AA.
AC P95867;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ALPHA-AMYLASE PRECURSOR (MALTO-OLIGOSYLITREHALOSE TREHALOHYDROLASE)
DE (TREZ) (EC 3.2.1.141).
GN TREZ.
OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
OX NCBI_TaxID=2287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35092 / DSM 1617 / P2;
RX MEDLINE=21332296; PubMed=11427726;
RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
RA Aways M.J., Chan-Weher C.C.-Y., Clausen I.G., Curtis B.A.,
RA De Moors A., Etraus G., Fletcher C., Gordon P.M.K.,
RA Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
RA Charlebois R.A., Doolittle W.F., Duguet M., Gaasterland T.,
RA Garrett R.L., Ragan M.A., Senses C.W., Van der Oost J.,
RA "The complete genome of the crenarchaeon Sulfolobus solfataricus P2."
RT Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
RL EMBL; Y08256; CNA69503.1; -.

DR EMBL; AE006815; AAK42272.1; -.
DR InterPro; IPR000461; Alpha_amylase.
DR ProSite; PS00128; alpha-amylase; 1.
DR Pfam; PF02922; isoamylase; 1.
KW Hydrolase; Glycosidase; Complete proteome.
SQ SEQUENCE 561 AA; 64370 MW; B00E403020F6B242 CRC64;

Query Match      60.4%; Score 1784.5; DB 17; Length 561;
Best Local Similarity 59.2%; Pred. No. 4.9e-111;
Matches 332; Conservative 90; Mismatches 130; Indels 9; Gaps 6;

Qy 3 SFGNIEKKNGIFKLPWAPYVNSVKLKLSPKLLPMEKNDEGFEVEIDDEENLTYSYII 61
Db 2 TFGYKLEDDGVTFLWAPYQKVKILNRGIYEMERDDKGYFTITLDNVVRGVRXYKIL 61

Qy 62 EDKREIPDPASRYQPLGVHDKSOLIRTDYQILDGKVKI--EDLIYELHVGTFSEOGNF 119
Db 62 DNSEVPDPASRYQPEGVHGYSEIISPDSEWDDNSVKKREDILVIELHIGTFTSEGT 121

Qy 120 KGVIEKLDYLKDLGITGIELMPVAQPPGNRDWYDGVFLYAVQNTYGGPWELAKLVNEAH 179
Db 122 EGVIRKUNYLKELGTAEIEMPIAQFFGKDKDWDGVTLYAVQNSYGGPSGFRKLVEAH 181

Qy 180 KRGIAVILDVYVNHIGPEGNYLLGLGPFSDRYKTPWGLTFNFDRCQDVRFKILENVE 239
Db 182 KGLAVILDVYVNHIGPEGNYLLGLGPFSEKYPKTPWGLTFNFDRCQDVRFKILENVE 241

Qy 240 YWTFKIDGLRLDAVHAIFDPSPHILQEIATKAHQLGKFKFVIAESDLNDPKIY--KDDC 297
Db 242 YWINEFVDDFRLDAVHAIFDPSPHILQEIATKAHQLGKFKFVIAESDLNDPKIY--KDDC 301

Qy 298 GYKIDAQWVDDFHAVHAFITKEKDYIYQDFGRIDIEKTFKDVYVYDGYKSYRGRTHG 357
Db 302 GYNDADQWVDDFHAVHAFITKEKDYIYQDFGRIDIEKTFKDVYVYDGYKSYRGRTHG 361

Qy 358 APVGDGDLPPRFVFIQNHQDVGNRNGERLSILTDKTYLMAATLYILSPYIPLIFMGEE 417
Db 362 KSVGDLGDLPPRFVFIQNHQDVGNRNGERLSILTDKTYLMAATLYILSPYIPLIFMGEE 421

Qy 418 YWTFKIDGLRLDAVHAIFDPSPHILQEIATKAHQLGKFKFVIAESDLNDPKIY--KDDC 477
Db 422 YGEENPFYIYDFSDPKLIQGVREGRRNGEQETDPSQDCTFNDKSKLSWKIDILSYFK 481

Qy 478 QLNIKRYNNCKRKEVREGNCITLIMEKIGIIASFDDIVINSKITGNLLI--GTGF 534
Db 482 SLIKIRKEYGLACNRKLSVNGVNYLVKNGCLAVYVFSKSVIEMKYSGLTVLSSNSP 541

Qy 535 PKLKKDELKLVNRGVGYQL 555
Db 542 PSQITESK-VELDKGFALYKL 561

RESULT 3
Q55088 PRELIMINARY; PRT; 559 AA.
AC Q55088;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ALPHA-AMYLASE.
OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
OX NCBI_TaxID=2287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KMI;
RA Kato M., Kettou M., Miura Y., Komeda T., Konishi Y., Shindo K.,
RA Kobayashi K., Iwamatsu A.;
RA "The gene analysis of the new amylases from the hyper thermophilic
RT archaea Sulfolobus";
RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.
```

```
DR EMBL; D641130; BAA1010.1; -.
DR InterPro; IPR000461; Alpha_amylase.
DR InterPro; IPR004193; isoamylase_N.
DR Pfam; PF00128; alpha-amylase; 1.
DR Pfam; PF02922; isoamylase_N; 1.
SQ SEQUENCE 559 AA; 64790 MW; 79FBE23A7CD38B4E CRC64;

Query Match 59.2%; Score 1748.5; DB 1; Length 559;
Best Local Similarity 58.9%; Pred. No. 1.2e-108;
Matches 329; Conservative 87; Mismatches 136; Indels 7; Gaps 5;

QY 3 SFGGTEKNKGIFKLPWVYNSVKLK-LSKKLIPMEKNDEGFEVEIDDEENLTYII 61
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2 TFYKIDGNEVFTLWAPYQKSVKLVLEKGLYEMERDDKGYFTITLNNVKVDKYYL 61
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 62 EDKRETPDASRQPLGVHDKSQLIRTDYQILDGLGVKIEDLIYELHVGTSQEQNFKG 121
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 62 DDASEIPDASRQPEGVHGPSQIIQESKEFNNETLKKEDLLIYELHVGTFPEGTFFG 121
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 122 VIEKLDYKDLGITGLMPVAQFPGNRDWGDVFLYAVQNTYGGPWELAKLVNEAHR 181
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 122 VIRKLDYKDLGITGLTAIEMPIAQFPCKRDWGDVFLYAVQNSYGGPEGRKLVDEAHR 181
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 182 GTAVLDVYNNHIGPGNYLLGLGPFSDRYKTPWGLTFNFDNRGCDQVRKFFILENVEY 241
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 182 GLGVILDVYNNHVGPGNYMVKLGPFYSQKYKTPWGLTFNFDNDAESDEVRKFFILENVEY 241
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 242 FKTEKIDGLRLDAVHAIFDNPSPKHILQETAERAKHQLGKFIASDLNDPKIV--KDCGY 299
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 242 IKEYNVGDFRLDAVHAIDTSPKHILEETADVVKHKNRIVIAESDLNDRVVPNPKCKGY 301
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 300 KIDAQWVDDFHAVHAFITKEDYDYQDFGRIEDIEKTEKDFVYDGYKSYRGRTHGAP 359
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 302 NIDAQWVDDFHSHIAYLGERQGYTDGFLDNDIVKSKDYVYDGYKSPRRKTHGAP 361
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 360 VGDLPKRFVFTQNDQVGNRGNRGERLSILDTKTYTLMAATLYILSPYPLIFMGEYY 419
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 362 VGELDCNFVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVY 421
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 420 ETNPFFESDFDPVLKGVREGRLKNNQMDPQSEEAFLSKLSWKIDEEVLDYKOL 479
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 422 EENPFYFSDFSKLIQVREGRLKNNQMDPQSEEAFLSKLSWKIDEEVLDYKOL 481
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 480 INIRKRYN-NCKRVKVRREGNCITLIMEKIGIIASFDIVINSKITGNLLIGI--GFPK 536
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 482 IKMRKLSIACDRRVVNGENWLLIKGREYSLYFSKSSIEVKYSGTLLSSNNSFPQ 541
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 537 KKKDELKLVNRGVGYQL 555
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 542 HIEEGK-YEFDRGFALYKL 559
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 4
Q9UWN9 PRELIMINARY; PRT; 559 AA.
AC Q9UWN9;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE MALTOOLIGOSYL TREHALOSE TREHALOHYDROLASE.
GN TREZ.
OS Sulfolobus shibatae.
OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
OX NCBI_TaxID=2286;
RN [1]
RP SEQUENCE FROM N.A.
RA Weiting Y., Wei C., Hui W., Li L., Cheng J.;
RT "Maltooligosyl trehalose trehalohydrolase from Sulfolobus shibatae.";
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF201335; AAF17553.1; -.
DR InterPro; IPR000461; Alpha_amylase.
DR InterPro; IPR004193; isoamylase_N.
```

```
DR Pfam; PF00128; alpha-amylase; 1.
DR Pfam; PF02922; isoamylase_N; 1.
KW Hydrolase.
SQ SEQUENCE 559 AA; 64619 MW; 8C0F0F38A5AE2846 CRC64;

Query Match 57.9%; Score 1711.5; DB 1; Length 559;
Best Local Similarity 58.0%; Pred. No. 3.6e-106;
Matches 324; Conservative 87; Mismatches 141; Indels 7; Gaps 5;

QY 3 SFGGTEKNKGIFKLPWVYNSVKLK-LSKKLIPMEKNDEGFEVEIDDEENLTYII 61
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2 TFYKIDGNEVFTLWAPYQKSVKLVLEKGLYEMERDDKGYFTITLNNVKVDKYYL 61
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 62 EDKRETPDASRQPLGVHDKSQLIRTDYQILDGLGVKIEDLIYELHVGTSQEQNFKG 121
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 62 DDASEIPDASRQPEGVHGPSQIIQESKEFNNETLKKEDLLIYELHVGTFPEGTFFG 121
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 122 VIEKLDYKDLGITGLMPVAQFPGNRDWGDVFLYAVQNTYGGPWELAKLVNEAHR 181
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 122 VIRKLDYKDLGITGLTAIEMPIAQFPCKRDWGDVFLYAVQNSYGGPEGRKLVDEAHR 181
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 182 GTAVLDVYNNHIGPGNYLLGLGPFSDRYKTPWGLTFNFDNRGCDQVRKFFILENVEY 241
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 182 GLGVILDVYNNHVGPGNYMVKLGPFYSQKYKTPWGLTFNFDNDAESDEVRKFFILENVEY 241
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 242 FKTEKIDGLRLDAVHAIFDNPSPKHILQETAERAKHQLGKFIASDLNDPKIV--KDCGY 299
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 242 IKEYNVGDFRLDAVHAIDTSPKHILEETADVVKHKNRIVIAESDLNDRVVPNPKCKGY 301
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 300 KIDAQWVDDFHAVHAFITKEDYDYQDFGRIEDIEKTEKDFVYDGYKSYRGRTHGAP 359
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 302 NIDAQWVDDFHSHIAYLGERQGYTDGFLDNDIVKSKDYVYDGYKSPRRKTHGAP 361
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 360 VGDLPKRFVFTQNDQVGNRGNRGERLSILDTKTYTLMAATLYILSPYPLIFMGEYY 419
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 362 VGELDCNFVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVY 421
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 420 ETNPFFESDFDPVLKGVREGRLKNNQMDPQSEEAFLSKLSWKIDEEVLDYKOL 479
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 422 EENPFYFSDFSKLIQVREGRLKNNQMDPQSEEAFLSKLSWKIDEEVLDYKOL 481
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 480 INIRKRYN-NCKRVKVRREGNCITLIMEKIGIIASFDIVINSKITGNLLIGI--GFPK 536
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 482 IKMRKLSIACDRRVVNGENWLLIKGREYSLYFSKSSIEVKYSGTLLSSNNSFPQ 541
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 537 KKKDELKLVNRGVGYQL 555
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 542 HIEEGK-YEFDRGFALYKL 559
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 5
Q973H4 PRELIMINARY; PRT; 359 AA.
AC Q973H4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE PUTATIVE ALPHA-AMYLASE.
GN ST0927.
OS Sulfolobus tokodaii.
OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
OX NCBI_TaxID=11955;
RN [1]
RP SEQUENCE FROM N.A.
RA Kavarabayasi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA Sekine M., Baba S.-I., Ankai A., Kosugi H., Hosoyama A., Fukui S.,
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
RA Oshima T., Kikuchi H.;
```


DR PROSITE: PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
KW Hydrolase; Complete proteome.
SQ SEQUENCE 583 AA; 65657 MW; 651CD7A918833E00 CRC64;

Query Match 29.3%; Score 865.5; DB 16; Length 583;
Best Local Similarity 38.2%; Pred. No. 1e-49; Mismatches 166; Indels 67; Gaps 20;

QY 4 FGGNIENK-KGIFKWPVNSVKLISK-KLIPMEKNDGFEFEIDIEENLTYSYII 61
DB 7 FGAQFGNGRTCEGLWAPDAREVRVETADGRWPLEGSDSGWEATL-PCPCTRYRYKI 65
QY 62 EDKREIPDPASYPQIPGVDHDKSOLIKTDYQIIDLGVK--KIE-----DIIIVELHVG 111
DB 66 DGRPGVPDPASQFQPDGVHGS-----QVLDDHGTYAMRVDEWRGPRWHEAVIYELHVG 118
QY 112 TFSQEGNFKGVTEKLDYKLDLGTITELMPVAQFPCGNRDWGYDGVFLYAVQNTYGGPWEL 171
DB 119 LF---GSYAEVERFLPRIVELGVTAVELMPLGFEFFGRNMGYDGLPFPAPASAYGTPEQL 175
QY 172 AKLVNEAHRGIAVLDVYVNHIGPEGNYLLG-LGPYFSDRYKTPWGLTFNFDGDCQV 230
DB 176 KHLIDSAGHGMGLVFDVYVNHIGPEGNYLLAQYAAAFRRDQTPWGQALDF-RRG--EV 232
QY 231 RKFLENVEYKTKIDGLRDVAHAIIDNSPKHILQEIABKAH-----QLGKFVIAES 285
DB 233 REFFYENALMWLLDYRVDRGLRDVAHAIIDNSPKHILQEIABKAH-----QLGKFVIAES 289
QY 286 DLNDPKIVKDDCGYKIDQAVDDDFHVAHAFITKEKDYQDFGR-IEDIEKTFKDFVY 344
DB 290 DNRASLQRQ--GY--DAQNDDGHHALVLLTGENDGYQDYPEPLRCLARCLAEFGFY 345
QY 345 DGKYSRYGRTHGAPVDLPPKFFVFIQNHQVGNRNGERLSILTDKTTVILMAATLYI 404
DB 346 QGEANRH-GRPGEPSADLAPAFVLFQNHQVGNRNGERLSVLAEPQALRLAQL 404
QY 405 LSPYIPLFMGEYETNPFPPFSDVPLIKGVREGRLKNNQ-----MTD 452
DB 405 LAPMIPLLFMGECAAREPFLYFTDQGE-LADAVREGRRKEFGFGRFEGATLASLPD 463
QY 453 POSEEAFLKSK-----LSWKIDEVDLYKOLINIRKY 486
DB 464 PNAVETFEERSPLGADCDPAWR-----GFYRQLLEIRHEH 498

RESULT 8
Q9RX51 PRELIMINARY; PRT; 600 AA.
AC Q9RX51;
DT 01-MAY-2000 (TremBLrel. 13, Created)
DT 01-MAY-2000 (TremBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TremBLrel. 19, Last annotation update)
DE MALTOOLIGOSYLTREHALOSE TREHALOXYDROLASE, PUTATIVE.
GN DR0464.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R1;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zaleski C.,
RA Makarov K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
radiodurans R1".
RL Science 286:1571-1577(1999).
DR EMBL; AE001905; AAF10042.1; -.

DR TIGR: DR0464; -.
DR InterPro; IPR000461; Alpha_amylase.
DR InterPro; IPR004193; isoamylase N.
DR Pfam; PF00128; alpha-amylase; 1.
DR Pfam; PF02922; isoamylase N; 1.
KW Hydrolase; Complete proteome.
SQ SEQUENCE 600 AA; 66909 MW; 594091EC093F8A44 CRC64;

Query Match 28.3%; Score 835.5; DB 16; Length 600;
Best Local Similarity 36.8%; Pred. No. 1.1e-47; Mismatches 191; Indels 31; Gaps 12;

QY 15 FKLWAPYVNSVKLISK-KLIPMEKNDGFEFEIDIEENLTYSYIEDKREIPDPASRY 74
DB 38 FRLWTSTARTVAVRVNGTETHVMTSLGGGIYELEL-PVPGARYLVL-DGVPTDPDYARF 95
QY 75 QPLGVHDKSOLI---RTDYQIIDLGVKVIKIEDLIIIVELHVGTSQEGNFKGVTEKLDYK 131
DB 96 LFDGVHGEAEVVDFTGTDADWHGKILADCVFEVHVHGTTPGTYRAAAEKLPYLKE 155
QY 132 LGITITELMPVAQFPCGNRDWGYDGVFLYAVQNTYGGPWELAKLVNEAHRGIAVLDVYV 191
DB 156 LGVTAIQVMPLAAPDQGRGWYDGAFAFYAPYGRPEDLMALVDAHRLGLGVFLDVY 215
QY 192 NHIGPEGNYLLGSP-YFSDRYKTPWGLTFNFDGDCQVQRKFIENVEYKTFKIDGL 250
DB 216 NHFGSGNYLSSVAPSYFTDFSSAWGMGLDYAE---PHMRRYVTGNARMWLRDYHFDGL 272
QY 251 RLDVAHAIIDNSPKHILQEIABKAHQLG--KFVIAESDLNDPKIVKDDCGYKIDQAVDD 308
DB 273 RLDATPYMTDDSETHLTELQAEIHELGGTHLLAEHRNLPDLVTN---HLDGIWTD 329
QY 309 FHAVHAFITTKDYVYQDF-GRIDIEKTFKDFVYDGVKYSRYGRTH--GAPVGDLP 365
DB 330 FHETRVLTGQEGYAGYRGAEALATIRGRWYEGQFWAVGEEHERGHPSDALEA 389
QY 366 RKFVFIQNHQVGNRNGERLSILTDKTTVILMAATLYILSPYIPLFMGEYETNPF 425
DB 390 PNFVYCIQNHQDIGNRPLGERLHQSDGVTLHEVRGAAALLPMTPLLFQGEWAASPTFQ 449
QY 426 FFSDFSDVPLIKGVREGRLKE-----NNQIDPQSEEAFLKSKLWKIDE-----BV 472
DB 450 FFSDHAGE-LQAVSEGRKKEFGFGFSGEDVDPQAEQTLNLSKLNWAEREGEHART 508
QY 473 LDYKOLINIRK 484
DB 509 LRLYRDLLRLR 520

RESULT 9
Q92U63 PRELIMINARY; PRT; 601 AA.
AC Q92U63;
DT 01-DEC-2001 (TremBLrel. 19, Created)
DT 01-DEC-2001 (TremBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TremBLrel. 19, Last annotation update)
DE PUTATIVE 1.4-ALPHA-GLUCAN BRANCHING ENZYME PROTEIN (EC
2.4.1.18).
OS Rhizobium meliloti (Sinorhizobium meliloti).
OG Plasmid pSymb (megaplasmid 2).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396508; PubMed=11481431;
RA Finan T.M., Weidner S., Wong K., Buhrmester J., Chain P.,
RA Vorhoeft F.J., Hernandez-Iucas I., Becker A., Cowie A., Gouzy J.,
RA Golding B., Puchler A.;
RT "The complete sequence of the 1,683-kb pSymb megaplasmid from the N2-
fixing endosymbiont Sinorhizobium meliloti".


```
Db 187 GSTADLKKLVDECHQIRIIMDGIYNHSEASSPLTQIDHDYWHHEPRDPDNN----- 240
QY 206 PYFSDRYKTPWGLTFNFD--DRGCD--QVRKFILENVEYWFKFKIDGLRLDAVHAIFDN 261
Db 241 -----WGPEFNTEHYDENLETYPARKFIDGTIVRYWGWETHLDGIRYDAARQIANY 290
QY 262 SPKH-ILQIEAKAHLGKFEVIAE-----SDLNDPKIVKDDCGYKIDAOQWDDPFHHA 312
Db 291 FEMHIAQEAKKTAGAKPFYVNAEHIPETTSIINLGP-----MDGCVHDSFYHT 340
QY 313 VHAFITKEKDYQDFGRIBEDIEKTKDFVYDYGKYSRYGRTHGAPVGLDPPRKFFVFI 372
Db 341 IKAHIC-----GDTFDLE-NLKDVI--DPKROGFLGATN-----VVNYL 376
QY 373 QNHG-----QVNGRNGERLSILTDK-----TTLMAATLYILSPVILPIEWGEY 418
Db 377 TNDHHDHIMVGLNR-----EFHDEAFRAKLGTAILMTAV-----GVPLIMWGEF 424
QY 419 YETNPFFFSDFSDPVLKIGVREGRLKENNMIDPOSEEAFKSKLSW-----KIDEEVL 473
Db 425 GEYKP-----KQDQ-----SKIDWTLLGNDLNRSLF 451
QY 474 DYYKQLINIRKRYNN 488
Db 452 DYHKGLIGLRK--NN 464

RESULT 12
Q93035 PRELIMINARY; PRT; 440 AA.
AC Q93035;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE BRANCHING ENZYME GLGB (FRAGMENT).
OS Myxococcus xanthus.
OC Bacteria; Proteobacteria; delta subdivision; Myxobacteria;
OC Myxococcales; Cystobacterineae; Myxococcaceae; Myxococcus.
OX NCBI_TaxID=34;
RN [1]
RP SEQUENCE FROM N.A.
RA Ueki T., Inouye S.;
RT "Identification of a new His-Asp phosphorelay signal transduction
RT system which regulates expression of a heat shock gene, lond, of
RT Myxococcus xanthus."
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF285783; AAK83002.1;
FT NON_TER 440 440
SQ SEQUENCE 440 AA; 48964 MW; 8D492EA5A4A92017 CRC64;
```

```
Query Match 14.2%; Score 420; DB 2; Length 440;
Best Local Similarity 32.0%; Pred. No. 3.7e-20;
Matches 133; Conservative 52; Mismatches 129; Indels 102; Gaps 17;

QY 52 EENLYTSYIIEKREIPDPASRYQPLGQVHDKSOLRTDYQLDLGKVKIEDLIYELHVG 111
Db 80 QENSTGSSIIYDHGEYWNAAQYSSPG-----FNEMIYYEMHVG 118
QY 112 TFSQE-----GNPKGVIEKLDYLDKLGITGIELMPVAQPPGNGRDWGYGVFLYAVONTYG 166
Db 119 TFHDSFGPGFGNNSAIARLDHVRDLGANMIKVMPAYEFAGDFSWGYNAAFPPAPESAYG 178
QY 167 GPWELAKLVNEAKHKGIAVILDVYVNHGP-----EGNYLLGLGYPFSD--RYKTPW 216
Db 179 HPNDMKRFYDEAHMRGIGVIFDVYVNHGYSFLPMMCFSGDCILGSGGEYFYNDWRKSTPW 238
QY 217 GLTFNFDRCQDQVRKFILENVEYWFKFKIDGLRLDAVH-----AIFD--NSP 263
Db 239 GDT--RPDGRPREVRIAYIDHSMNMLTSPFGGLRWDAIKYMTQNGSTDTALPDARVVF 296
QY 264 KHIQIEAKAHLGKFEVIAESDLNDPKIVKDDC-----GYKIDAQWDDFHHAVHAFI 317
```

```
Db 297 RSINREI--NATPWKISIAE--DFGGGDFITNDATSDTSGAGFDSQWGGDFVHAIRAAV 353
QY 318 TREKDYIYQDFGRIEDIEKTKDFVYDYGKYSRYGRTHGAPVGLDPPRFVFIQNHQ 377
Db 354 IAS-----NDSGR--DMNS-----VRNATQRYSGR-HTAR-----VIYSESHDE 390
QY 378 V-----GNRNGERLSILTDKTYLMAATLYILSPVILPIEWGEY 420
Db 391 VANGKARVPEEINPGNAGS-----WNAKARSLAAGVETSPGIPMIFOGQEFLE 440

RESULT 13
Q59242 PRELIMINARY; PRT; 652 AA.
AC Q59242; O08495;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE 1,4-ALPHA-GLUCAN BRANCHING ENZYME (EC 2.4.1.18)
DE (GLYCOGEN BRANCHING ENZYME) (1,4-ALPHA-GLUCAN BRANCHING ENZYME)
DE (AMYLO-(1,4 TO 1,6)TRANSGLUCOSIDASE)
DE (AMYLO-(1,4-1,6)-TRANSGLYCOSYLASE).
GN GLGB.
OS Bacillus stearothermophilus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Geobacillus.
OX NCBI_TaxID=1422;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TRBE14;
RX MEDLINE=95031021; PubMed=7944355;
RA Takata H., Takaha T., Kuriki T., Okada S., Takagi M., Imanaka T.;
RT "Properties and active center of the thermostable branching enzyme
RT from Bacillus stearothermophilus."
RL Appl. Environ. Microbiol. 60:3096-3104(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=TRBE14;
RA Takata H.;
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=TRBE14;
RA Takata H., Takaha T., Okada S., Takagi M., Imanaka T.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: FORMATION OF 1,6-GLUCOSIDIC LINKAGES OF
CC GLYCOGEN.
CC -1- PATHWAY: THIRD STEP IN GLYCOGEN BIOSYNTHESIS.
CC -1- MISCELLANEOUS: OPTIMAL ACTIVITY AT APPROXIMATELY 39 DEGREES
CC CELSIUS.
CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
DR EMBL; D87026; BAA19588.1;
DR InterPro; IPR000461; Alpha_amylase.
DR InterPro; IPR004193; isoamylase_N.
DR Pfam; PF00128; alpha-amylase; 1.
DR Pfam; PF02922; isoamylase_N; 1.
KW Glycogen biosynthesis; transferase; Glycosyltransferase.
FT ACT_SITE 308 308 BY SIMILARITY.
FT ACT_SITE 351 351 BY SIMILARITY.
FT ACT_SITE 419 419 BY SIMILARITY.
SQ SEQUENCE 652 AA; 76793 MW; 4591EB414A0E3FEF CRC64;
```

```
Query Match 12.2%; Score 360; DB 2; Length 652;
Best Local Similarity 25.3%; Pred. No. 6.3e-16;
Matches 142; Conservative 72; Mismatches 186; Indels 162; Gaps 25;

QY 4 FGNIENKNG-----IFKLWAPYVNSVKL-----KLSKLIPEKNDGFEFE 47
Db 26 FGAHVINEGGKVGTRFCVWAPHAREVLVGSFNDWGDGDFRLEK-----VNDEGVMTIV 79
```

```
Qy 48 IDDIENITYSY-IIDKKREI---PDPASRYQPLGVHDKSOLIRTD---YOILDIG----- 96
Db 80 VPENLEGLHYKYEIVTPDQVLFADPYAFYSELPHPTAS---IAYDLKGYOWNQDSWKRK 137
Qy 97 ---KVKIEDLIIYELHVGTF--SQEGNEKGVIEKLD-----YKDLGITGIELMPVAQPPG 147
Db 138 KRRRIYDQPMVYIELHFGSKKDGREYTYREMADELISYVDLHGFTHEILLPLVEHPL 197
Qy 148 NRDMGYDGVFLYAVONTYGGPWELAKLNEAHKRGIAVILDVYNNHI--GPEGNYLLGLG 205
Db 198 DRSGYQGTGYAYTSRGTGTHDPMYFVDRCHQAGIGVMDWPCGHCKDAHGLYMDGA 257
Qy 206 PYF-----SDRYKTPMGUTGTFNDRGCDQVRKFFILENVYWFKTFKIDGLRLDAY----- 255
Db 258 PTYEYANEKDRENYWG-TANF-DLQKPEVRSFLISNALFWLEYHYHIDGFRVDVANMLY 315
Qy 256 ---HAIFDNS-PKHILQIEAEK--AHQLGKFVTAESDLNDPKIVKD-----DCGYKID 302
Db 316 WPNDRLYENPYAVEFLKRLKNEAVFADPNALMIAEDSTDWPKVTAPTTEGGGLGFNYKWN 375
Qy 303 AQWYDDF-----HHAHVAFITTEKDYQYQDFGRIEDIEKTFKDVYVDGKYSRYR 352
Db 376 MGWMDMLKYMETPPYERRHHVHQVTSLLYAYS-----ENFILPFSHDEVVHGKKS--- 427
Qy 353 GRTHGAPVDLPPRKFFVFIQNHQVGNRNG-----ERLSILTDKTYTLMAATLYILSPYI 409
Db 428 -----LLNKMPGSYEERKFAQLRLLYGYMA-----HPGK 456
Qy 410 PLIFMGEYETNPPFFSDFSDPVLKGVREGLRKNNQMDPQSEAFLLKSLW--- 466
Db 457 KLLFMGNEF-----AQDFEWDKFEDELDDWLVF 482
Qy 467 --KIDEEVLDYKQLINIRKRY 486
Db 483 DFLHRKNDYMKELIACYKRY 504
RESULT 14
P73608 ID P73608 PRELIMINARY; PRT; 707 AA.
AC P73608:
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE GLYCEN OPERON PROTEIN GLGX.
GN GLGX OR SLR1857.
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hiroseawa M., Sugitara M., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA Shimpo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1996).
DR EMBL; D90908; BAA17652.1; -.
DR HSPG; P10342; 1BF2.
DR InterPro; IPR000461; Alpha_amylase.
DR InterPro; IPR004193; isoamylase-N.
DR Pfam; PF00128; alpha-amylase-1.
DR Pfam; PF02922; isoamylase-N; 1.
KW Complete proteome.
SQ SEQUENCE 707 AA; 79895 MW; CFA27B3C86318DB1 CRC64;
```

Query Match 11.38; Score 333.5; DB 16; Length 707;
Best Local Similarity 23.18; Pred. No. 4.1e-14;

```
Matches 146; Conservative 83; Mismatches 197; Indels 207; Gaps 29;
Qy 2 FSGGNTKNGKIFKLWAPYVNSVKLSKSLIPMEKNDEGFEVE----- 47
Db 23 FFGATIVPGGVNSIYSSHSHTACTLVLFEKRAP-----QPFVEIPPESEFRIGNVYCMV 77
Qy 48 IDDIENITYSY-IIDKKREI---PDPASRYQPLGVHDKSOLIRTDYQILDIGK----- 97
Db 78 VFDLDFENLEYGYRMEG-----PNNFQOQHWFDPKSVLLDPYAKVVSGRDVGWGTQDNW 130
Qy 98 -----VKIEDLIIYELHVGTFSQE-----GNFKGVI 123
Db 131 DDIYQHGRGLSFFDDWENDSPLDVPLEDMVYIEMHVRGFTKDPSSGYKENHRTGAGIL 190
Qy 124 EKLDYLDLGTGTIELMPVAQF-----PGNRD-----WGYDGVFLYAVONTYGGPW 169
Db 191 SKIPLYLOELGVNTIELMPIFEDEFESHSRYHPETGEFLVNWYGVSTVNFPAKAGYATG 250
Qy 170 -----ELAKLVNEAHKRGIAVILDVYNNHIGPEGNYLLGLGYPFSDR-----YKT 214
Db 251 KFGWQIDELKNLVKELHKGISVILDVYVFNHTA-EGN---ERGPTISFRGLDNKTYMLT 306
Qy 215 PWGLTFNFDGRG---CDQ---VRKFILENVYWFKTFKIDGLRLDAYHAIIDNSPKHILQ 268
Db 307 PEGYFNFSGTGNLNCNNPIVRGMVLDCLRYTAEFHIDGFRFD-LASILGRDP----- 360
Qy 269 EIAEKAHOLGRFVIAESDLNDP---KIVKD---DCG--YKID-----AQMVDFFH 311
Db 361 ---WGYPPLANPPILETALFDPILARSLKLAENDAGLVQVSGFSPYGRWAEWNGKYRD 416
Qy 312 AVHAFITTEKDY-----YYQDFGRIEDIEKTF-----KDFVYVDGKYS 349
Db 417 TVRKFIKGDAGVICEMAORLQSGPDLYOGAGRPSTSFNFVTAHGDFTLADLVAYNGKH 476
Qy 350 RYRGTHGAPVDLPPRKFFVFIQNHQD-OVCNRCNGERLSILTDKTYLM-AATLYILSP 407
Db 477 YANGENGNDGAND-----NYSWNCVGEPTDNPDLRLRARQMRNAIALLVSQ 525
Qy 408 YIPLIFMGE-----YYETNPPFFSDFSDPVLKGVREGLRKNNQMDPQSE 456
Db 526 GVPMLMGDEMGKTDGNNNTYCHDSFENLWLN-----WHLEQN-----QAW 567
Qy 457 EAFLLKSLKSKIDEVLDYKQLINIRKRYNNC 489
Db 568 FRFVKHCIAFRLAHPVL-----RNSEHFQNC 593
RESULT 15
Q97S07 ID Q97S07 PRELIMINARY; PRT; 1280 AA.
AC Q97S07:
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ALKALINE AMYLOPULLANASE, PUTATIVE.
GN SP0268.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=TI6R4;
RX MEDLINE=21357209; PubMed=11463916;
RA Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
RA Peterson S., Heidelberg J., DeBoy R.T., Haft D.H., Dodson R.J.,
RA Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
RA Unayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
RA Holtzapple E., Khouri H., Wolf A.M., Utterback T.R., Hansen C.L.,
RA McDonald L.A., Feldblyum T.V., Angiuoli S., Dickenson T., Hickey E.K.,
RA Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RT "Complete genome sequence of a virulent isolate of Streptococcus
```

Search completed: July 15, 2002, 12:24:10
Job time: 392 sec

